

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:15:10 ; Search time 1741.97 Seconds
(without alignments)
10198.307 Million cell updates/sec

Title: US-10-077-130-6_COPY_19500_22500
Perfect score: 3001
Sequence: 1 cctacggagcagaactcggg.....ccacctggagcaagaacgga 3001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

- 1: Geneseqn1980s.*
- 2: Geneseqn1990s.*
- 3: Geneseqn2000s.*
- 4: Geneseqn2001as.*
- 5: Geneseqn2001bs.*
- 6: Geneseqn2002as.*
- 7: Geneseqn2002bs.*
- 8: Geneseqn2003as.*
- 9: Geneseqn2003bs.*
- 10: Geneseqn2003cs.*
- 11: Geneseqn2003ds.*
- 12: Geneseqn2004as.*
- 13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3001	100.0	4936	AAH46904	Aah46904 CDNA enco
2	3001	100.0	8106	ABX11641	Abx11641 Human ser
3	3001	100.0	24120	ABX11642	Abx11642 Human ser
4	2999.4	99.9	5007	AAc62286	AAc62286 CDNA enco
5	2999.4	99.9	7928	AAc62287	AAc62287 CDNA enco
6	2997.8	99.9	5207	AAI43909	AAI43909 Human kin
7	2997.8	99.9	5207	AAI43908	AAI43908 Human kin
8	2996.2	99.8	5454	AAc38865	AAc38865 Human kin
9	2643.4	88.1	4175	AAc62285	AAc62285 CDNA enco
10	1706.8	56.9	3225	AAc26467	AAc26467 Human kin
11	223.2	7.4	7710	AAf44662	AAf44662 Novel pro
12	223.2	7.4	7710	ADI29360	ADI29360 Human MAR
13	223.2	7.4	7789	AAc30565	AAc30565 Human kin
14	223.2	7.4	9698	AAE47675	AAE47675 Human NOV
15	223.2	7.4	9698	ADJ78945	ADJ78945 Human NOV
16	223.2	7.4	9807	ADB79958	ADB79958 Human kin
17	223.2	7.4	10662	ADC99127	ADC99127 Human KPP
18	204	6.8	9930	ADe47671	ADe47671 Human NOV
19	204	6.8	9930	ADJ78941	ADJ78941 Human NOV
20	204	6.8	10122	ADe47673	ADe47673 Human NOV

21	204	6.8	10122	12	ADJ78943	Adj78943 Human NOV
22	173.6	5.8	860	10	ADe47677	ADe47677 Human NOV
23	173.6	5.8	860	12	ADJ78947	ADJ78947 Human NOV
24	172	5.7	1429	2	AAx34657	AAx34657 Murine ZI
25	170.2	5.7	1365	2	ADo71726	ADo71726 Human gen
26	170.2	5.7	2055	13	ADN60271	ADN60271 Human ZIP
27	170.2	5.7	2055	13	ADR83403	ADR83403 Human ZIP
28	170.2	5.7	2079	4	AAK94258	AAK94258 Human ful
29	170.2	5.7	2079	12	ADL30841	ADL30841 Full leng
30	170.2	5.7	2079	13	ADQ68012	ADQ68012 Recombina
31	170.2	5.7	2104	12	ADQ84258	ADQ84258 Human tum
32	170.2	5.7	2104	13	ADQ85644	ADQ85644 Human tum
33	170.2	5.7	2104	13	ADQ86740	ADQ86740 Human tum
34	170.2	5.7	2226	10	ADD29770	ADD29770 Human tum
35	168.6	5.6	2105	6	ABV72291	ABV72291 Nucleotid
36	168.6	5.6	2105	13	ADN60267	ADN60267 Human dea
37	168.6	5.6	2105	13	ADN60269	ADN60269 Human ZIP
38	168.6	5.6	2105	13	ADR83493	ADR83493 Human dea
39	168.6	5.6	2132	2	AAx34656	AAx34656 Human ZIP
40	168.6	5.6	2132	8	ACA90235	ACA90235 Deatch as
41	168.6	5.6	2132	12	ADO09424	ADO09424 Novel hum
42	168.6	5.6	2224	4	AAH16158	AAH16158 Human CDN
43	168.6	5.6	2224	5	AAH78068	AAH78068 Nucleotid
44	164	5.5	1514	10	ADB58463	ADB58463 Toxicity-
45	164	5.5	1514	10	ADB53052	ADB53052 Primary r

ALIGNMENTS

RESULT 1						
AAH46904						
ID	AAH46904	standard; cDNA; 4936 BP.				
XX						
AC	AAH46904;					
XX						
DT	25-SEP-2001	(first entry)				
XX						
DE	CDNA encoding human protein kinase SGK145.					
XX						
KW	Protein kinase; enzyme; cytostatic; nootropic; neuroprotective; human;					
KW	antiparkinsonian; virucide; antibacterial; antifungal; antimigraine;					
KW	analgesic; hypotensive; hypertensive; immunosuppressive; anxiolytic;					
KW	antipruritic; antirheumatic; antiarthritic; ophthalmological; anorectic;					
KW	osteopathic; thrombolytic; antiarteriosclerotic; antiasthmatic;					
KW	vasotropic; antidiabetic; gene therapy; ss.					
XX						
OS	Homo sapiens.					
XX						
FH	Key	Location/Qualifiers				
CDS		1..4851				
FT		/*tag= a				
XX						
PN	WO200155356-A2.					
XX						
PD	02-AUG-2001.					
XX						
PF	25-JAN-2001; 2001WO-US002337.					
XX						
PR	25-JAN-2000; 2000US-0178078P.					
PR	31-JAN-2000; 2000US-0179384P.					
PR	17-FEB-2000; 2000US-0183173P.					
PR	17-MAR-2000; 2000US-0190162P.					
PR	29-MAR-2000; 2000US-0191404P.					
PR	13-NOV-2000; 2000US-0247013P.					
XX						
PA	(SUGE-) SUGEN INC.					
PI	Plowman G, Whyte D, Manning G, Sudarsanam S, Martinez R;					
XX						
DR	WPI; 2001-476202/51.					
DR	P-PSDB; AAB85504.					
XX						

PT Kinase polypeptides useful for treating cancers, Alzheimer's disease,
PT viral infections, diabetes, obesity, organ transplant rejection and
PT rheumatoid arthritis.

Example 1; Page 204-205; 218pp; English.

The invention provides human protein kinases and protein kinase-like enzymes and polynucleotides encoding the polypeptides. The kinase polypeptides and their modulators are useful for treating a disease or disorder such as cancer, immune-related diseases, cardiovascular disease, brain or neuronal-associated disease and metabolic disorders, including cancers of tissues, cancers of hematopoietic origin, diseases of the central nervous system, diseases of the peripheral nervous system, Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic lateral sclerosis, viral infections, infections caused by prions, bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, hypertension, psychotic disorders, neurological disorders, dyskinesias, metabolic disorders, and organ transplant rejection. They are also useful for treating rhinitis, autoimmunity, atherosclerosis, psoriasis, osteoarthritis, asthma, chronic inflammatory pelvic disease, chronic inflammatory bowel disease, rheumatoid arthritis, metabolic disorders such as diabetes, obesity, cardiovascular diseases such as reperfusion injury, coronary thrombosis, clotting disorders and atherosclerosis, ocular diseases such as glaucoma, retinopathy and macular degeneration, psychiatric and neurological disorders such as anxiety, schizophrenia, dementia, manic depression, etc. The polynucleotides are useful in gene therapy techniques to treat the above mentioned disorders. Sequences AAH46891-46922 represent human protein kinases encoding cDNA molecules

SQ Sequence 4936 BP; 923 A; 1695 C; 1524 G; 794 T; 0 U; 0 Other;

Query Match 100.0%; Score 3001; DB 4; Length 4936;

Best Local Similarity 100.0%; Pred. No. 0

Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1	CTTACGGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTGCGCCGCGCTGAG	60
450	CTTACGGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTGCGCCGCGCTGAG	509
61	CCACCCGCTGTGTCACGGGCGTCTGGACCGAGCTTTTGAGACCGCGAAGACCTCATCTCTCAT	120
510	CCACCCGCTGTGTCACGGGCGTCTGGACCGAGCTTTTGAGACCGCGAAGACCTCATCTCTCAT	569
121	CTTGGAGCTGTGCTCATCCGAGGAGCTGCTGGACCGCTGTACAGGAAGGGCGTGGTGAC	180
570	CTTGGAGCTGTGCTCATCCGAGGAGCTGCTGGACCGCTGTACAGGAAGGGCGTGGTGAC	629
181	GGAGGCGGAGGTCAGAGTCTACATCCAGCAGCTGGTGGAGGGGTGCATCACTGCAACAG	240
630	GGAGGCGGAGGTCAGAGTCTACATCCAGCAGCTGGTGGAGGGGTGCATCACTGCAACAG	689
241	CCATGGCGTTCTCCACCTGGGCATTAAGCCCTCTAAACATCTGATGGTGCATCCTCGCCG	300
690	CCATGGCGTTCTCCACCTGGGCATTAAGCCCTCTAAACATCTGATGGTGCATCCTCGCCG	749
301	GGAGAGCATTTAAATCTCGCACTTTGGGCTTTGCCAGAACATCACCCAGCAGAGTGCA	360
750	GGAGAGCATTTAAATCTCGCACTTTGGGCTTTGCCAGAACATCACCCAGCAGAGTGCA	809
361	GTTCAGCCAGTACGGCTCCCTCGATGTCGTCTCCCGGAGATCATCCAGCAGAACCTGT	420
810	GTTCAGCCAGTACGGCTCCCTCGATGTCGTCTCCCGGAGATCATCCAGCAGAACCTGT	869
421	GAGCGAAGCCTCCGACATTTGGGCCATGGGTGTCATCTCTTAAGCTGACCTGTCTC	480
870	GAGCGAAGCCTCCGACATTTGGGCCATGGGTGTCATCTCTTAAGCTGACCTGTCTC	929
481	ATCCCCATTTGCGCGGAGGTGACCGTGCCACCCCTCTGAAAGTCTCTGGAGGGCGCT	540
930	ATCCCCATTTGCGCGGAGGTGACCGTGCCACCCCTCTGAAAGTCTCTGGAGGGCGCT	989
541	GTCTAGGAGCAGCCCATCGGTGCGCCACTCAGCGAAGAGCGCCAAAGACTTATCAAGCG	600

Db 2070 TTCTGCCCCCCAGGAGGGCTGCAGCCCCCACCAGCAGTTGCCCCATGCCCTCTGGCTC 2129
Qy 1681 CTTCCCTTCAGGATCTTGCAAGAGGCCCCCTTAGTACCTCTCAAGCCCTTCTTGGGACA 1740
Db 2130 CTTCCCTTCAGGATCTTGCAAGAGGCCCCCTTAGTACCTCTCAAGCCCTTCTTGGGACA 2189
Qy 1741 GCCCCAGGACCCCTGCCCCCTGCAAGCAAGAGCCCCCATTTGGACTCTAGATGGGGCC 1800
Db 2190 GCCCCAGGACCCCTGCCCCCTGCAAGCAAGAGCCCCCATTTGGACTCTAGATGGGGCC 2249
Qy 1801 TGGAGACATCTCTCTTCTCTGGAGGCCCCAAACCCGGGCCCTTGCAAGTTCCCCAGAGGTGAGC 1860
Db 2250 TGGAGACATCTCTCTTCTCTGGAGGCCCCAAACCCGGGCCCTTGCAAGTTCCCCAGAGGTGAGC 2309
Qy 1861 CTCCAGGCGAGCTCTTCCCAAGTAGCTCCCTCAGAGGTGGGCTCTCTCCAGGTGGGAC 1920
Db 2310 CTCCAGGCGAGCTCTTCCCAAGTAGCTCCCTCAGAGGTGGGCTCTCTCCAGGTGGGAC 2369
Qy 1921 AGAGCCTGGCCCCCTCTGGATGGAGGGCTGGACCCAGGAGCTGAGGATCTGCCGA 1980
Db 2370 AGAGCCTGGCCCCCTCTGGATGGAGGGCTGGACCCAGGAGCTGAGGATCTGCCGA 2429
Qy 1981 CTCCACACCCACCTTGACGCGGCTCTCAGAAACAGGCGACCATGGCGAAGTTCTCCCTGGG 2040
Db 2430 CTCCACACCCACCTTGACGCGGCTCTCAGAAACAGGCGACCATGGCGAAGTTCTCCCTGGG 2489
Qy 2041 TGGTCGGGGGGCTACGAGCGGCTGGCTGTGTATGGACCTTTTGGCTTTGTGTGAGATGC 2100
Db 2490 TGGTCGGGGGGCTACGAGCGGCTGGCTGTGTATGGACCTTTTGGCTTTGTGTGAGATGC 2549
Qy 2101 AGGGGGATGCTGGGGCAGGGGCCATGTGGCCAGGATAGCTGGGGCTGTGCCAGTC 2160
Db 2550 AGGGGGATGCTGGGGCAGGGGCCATGTGGCCAGGATAGCTGGGGCTGTGCCAGTC 2609
Qy 2161 GGAGGAGGAGGAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGGAGCAGCAGGAGGC 2220
Db 2610 GGAGGAGGAGGAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGGAGCAGCAGGAGGC 2669
Qy 2221 CAGGGCTGAGAGCCACTGCCCCAGGTGAGTGCAGAGGCTGTGTGAGGTGCGAGGCG 2280
Db 2670 CAGGGCTGAGAGCCACTGCCCCAGGTGAGTGCAGAGGCTGTGTGAGGTGCGAGGCG 2729
Qy 2281 TCCACACGAGGAGCTCCAGAGCCACCCCATGGAGGAGCATCGGGCAGGTCTCCCTGGT 2340
Db 2730 TCCACACGAGGAGCTCCAGAGCCACCCCATGGAGGAGCATCGGGCAGGTCTCCCTGGT 2789
Qy 2341 GCAGATCCGGACCTGTGAGGTGATGCGAGGCGCGGACACAATATCCCTGGACATTTTC 2400
Db 2790 GCAGATCCGGACCTGTGAGGTGATGCGAGGCGCGGACACAATATCCCTGGACATTTTC 2849
Qy 2401 CGAGGTGGAACCCCGCTACCTCAACCTCTCAGACCTGTATCGATATCAAGTACCTCCCAT 2460
Db 2850 CGAGGTGGAACCCCGCTACCTCAACCTCTCAGACCTGTATCGATATCAAGTACCTCCCAT 2909
Qy 2461 CGAGTTTATGATCTTCAGGAAGTCCCAAGTCCGCTCAGCGAGCGCCCTCCCCCAT 2520
Db 2910 CGAGTTTATGATCTTCAGGAAGTCCCAAGTCCGCTCAGCGAGCGCCCTCCCCCAT 2969
Qy 2521 GGCTGAGGAGGAGTGGCCGAGTTCCCGAGGCCACAGTGGCCCTGGCCAGGTGAACCTGGG 2580
Db 2970 GGCTGAGGAGGAGTGGCCGAGTTCCCGAGGCCACAGTGGCCCTGGCCAGGTGAACCTGGG 3029
Qy 2581 CCCCCACGAGGCTGGAGATCACAGAGGAGTGCAGAGATGTGACGCGCTGTCTGGCAGA 2640
Db 3030 CCCCCACGAGGCTGGAGATCACAGAGGAGTGCAGAGATGTGACGCGCTGTCTGGCAGA 3089
Qy 2641 GGCTGCCGTGGGAGGAGCGCAGTGGTCTCTCGCCGTCAAGCGCTCTTCCACTTCCC 2700
Db 3090 GGCTGCCGTGGGAGGAGCGCAGTGGTCTCTCGCCGTCAAGCGCTCTTCCACTTCCC 3149
Qy 2701 TGGAGGCGACCTGCCGTGGATGAGCTGCGAGAGTGGGGCTGCGTGAAGAGTGAAGGC 2760
Db 3150 TGGAGGCGACCTGCCGTGGATGAGCTGCGAGAGTGGGGCTGCGTGAAGAGTGAAGGC 3209

Qy 2761 CTCCTGGAGCACATCTCCCGGATCTCGAAGGGCAGGCCGGAAGGTTCTGGAGAAGAGGG 2820
Db 3210 CTCCTGGAGCACATCTCCCGGATCTCGAAGGGCAGGCCGGAAGGTTCTGGAGAAGAGGG 3269
Qy 2821 GCCCCCAGAGAAAGCAGGSCCTTGTCTTCCGGCTCTCAGGTCTGAAGAGCTGGGA 2880
Db 3270 GCCCCCAGAGAAAGCAGGSCCTTGTCTTCCGGCTCTCAGGTCTGAAGAGCTGGGA 3329
Qy 2881 CCGAGCCCGACATCTCTAAGGGAGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 2940
Db 3330 CCGAGCCCGACATCTCTAAGGGAGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 3389
Qy 2941 GACACTGGCTGCGCAGGTGTGAGCCAGCCAGCTGCCCCAGCCACCTTGGAGCAAGACGG 3000
Db 3390 GACACTGGCTGCGCAGGTGTGAGCCAGCCAGCTGCCCCAGCCACCTTGGAGCAAGACGG 3449
Qy 3001 A 3001
Db 3450 A 3450

RESULT 2
ABX11641
ID ABX11641 standard; cDNA; 8106 BP.
XX
AC ABX11641;
XX
DT 09-MAY-2003 (first entry)
XX
DE Human serine/threonine or protein kinase 59079, cDNA.
XX
KW Human; ss; gene; serine/threonine kinase; protein kinase; 59079;
cardiovascular disease; heart failure; myocardial infarction;
blood vessel disorder; atherosclerosis; Kaposi's sarcoma;
blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
haemolytic anaemia; cellular proliferative disorder; cancer;
protein kinase disorder; autoimmune disorder; diabetes mellitus;
psoriasis; inflammatory bowel disease; rheumatoid arthritis;
multiple sclerosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..71
FT CDS /*tag= a
FT 72..7964
FT /*tag= b
FT /*product= "Kinase 59079"
FT /*note= "This CDS is specifically claimed in claim 2"
FT 3'UTR 7965..8106
FT /*tag= c
XX
PN US2002168742-A1.
XX
PD 14-NOV-2002.
XX
PF 15-FEB-2002; 2002US-00077130.
XX
PR 15-FEB-2001; 2001US-0269201P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Kapeller-Libermann R, Acton SL;
XX
PI WPI; 2003-298729/29.
XX
DR P-PSDB; ABG76186.
XX
PT Novel isolated human protein kinase, designated 59079 or 12599
polypeptide, useful as diagnostic and therapeutic agents for preventing
cardiovascular diseases, proliferative disorders, and protein kinase
disorders.
XX

PS	Claim 2; Page 39-48; 119pp; English.	
XX	The invention relates to an isolated human serine/threonine or protein kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule comprising at least 85% identity to the nucleic acids appearing as ABX11641 and ABX11642 or their complement, a naturally occurring variant of the kinases or their fragments. Also included are a non-human host cell containing the nucleic acids, an antibody specific for the proteins, identifying a compound which binds to the kinase (by contacting the kinase or a cell expressing the kinase with a test compound and determining whether the kinase binds to the test compound) and modulating the activity of kinase using the identified compound. The kinases and their encoding nucleic acids are useful as diagnostic and therapeutic agents for preventing a disease or condition associated with an aberrant or unwanted 59079 or 12599 activity in a subject, including cardiovascular diseases such as heart failure, and myocardial infarction; disorders involving blood vessels such as atherosclerosis, and Kaposi's sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia, Hodgkin's disease, hemolytic anaemia; cellular proliferative disorders such as cancer; and protein kinase disorders such as autoimmune disorders, diabetes mellitus, psoriasis, inflammatory bowel disease, rheumatoid arthritis, and multiple sclerosis (many examples of diseases and disorders are included in the specification). The kinases, their encoding nucleic acids and antibodies are useful in screening assays, detection assays (e.g. forensic biology), and predictive medicine (e.g. diagnostic assays, prognostic assays, and monitoring clinical trials and pharmacogenomics). The kinases and their encoding nucleic acids are useful as query sequences to perform a search against public databases to identify other family members or related sequences. The present sequence encodes the kinase 59079	
XX	Sequence 8106 BP; 1603 A; 2684 C; 2530 G; 1289 T; 0 U; 0 Other;	
SQ	Query Match 100.0%; Score 3001; DB 8; Length 8106; Best Local Similarity 100.0%; Pred. No. 0; Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 CCTACGGAGCAGAACTCGGGCCCAAGGCATACAGGAGGGAGACATCTTGGCCGCGCTGAG 60	541 GTCTATGAGCAGAGCCCATGGCTGCCACCTCAGCGAAGACGCCAAAGACTTTCATCAAGGC 600
DB	3557 CCTACGGAGCAGAACTCGGGCCCAAGGCATACAGGAGGGAGACATCTTGGCCGCGCTGAG 3616	4097 GTCTATGAGCAGAGCCCATGGCTGCCACCTCAGCGAAGACGCCAAAGACTTTCATCAAGGC 4156
QY	61 CCACCCGCTGTCTACGGGGCTGTCTGGACCAAGTTTGAGACCGCAAGACCTTCATCTCAT 120	601 TACGCTGACAGAGCCCTCTCAGGGCCCGGCTAGTGGCGGCCAGTGCTCTCTCCCAACCCCTG 660
DB	3617 CCACCCGCTGTCTACGGGGCTGTCTGGACCAAGTTTGAGACCGCAAGACCTTCATCTCAT 3676	4157 TACGCTGACAGAGCCCTCTCAGGGCCCGGCTAGTGGCGGCCAGTGCTCTCTCCCAACCCCTG 4216
QY	121 CTGAGAGTGTGCTATCCGAGGAGCTGTGTGACCGCTGTACAGGAAGGGCGTGTGAC 180	661 GTTCTGAAATTCATGCTGCGGAGGAGGCCACTTTCATCAACCAAGCAGCTCAAGTT 720
DB	3677 CTTGAGAGTGTGCTATCCGAGGAGCTGTGTGACCGCTGTACAGGAAGGGCGTGTGAC 3736	4217 GTTCTGAAATTCATGCTGCGGAGGAGGCCACTTTCATCAACCAAGCAGCTCAAGTT 4276
QY	181 GGAGGCCGAGGTCAAGGTCTACATCCAGCAGCTGTGGAGGGGCTGCACCTACCTGCACG 240	721 CCTCTGGCCCGCAAGTCTGTCGAGAGCTTCCCTGATGAGCTACAGTCCATCTCTGGTAT 780
DB	3737 GGAGGCCGAGGTCAAGGTCTACATCCAGCAGCTGTGGAGGGGCTGCACCTACCTGCACG 240	4277 CCTCTGGCCCGCAAGTCTGTCGAGAGCTTCCCTGATGAGCTACAGTCCATCTCTGGTAT 4336
QY	241 CCAATGGCGTTCTCCACTGGACATAAAGCCCTCTAAATCCTGATGGTGCATCCTGCCCG 300	781 GCGCTCCATCCCTGAGCTGTCTGGGGGCCCAACCGACAGCCCTCCCTCGGGCTAGCCCCG 840
DB	3797 CCATGGCGTTCTCCACTGGACATAAAGCCCTCTAAATCCTGATGGTGCATCCTGCCCG 3856	4337 GCGCTCCATCCCTGAGCTGTCTGGGGGCCCAACCGACAGCCCTCCCTCGGGCTAGCCCCG 4396
QY	301 GGAAGACATTAATAATTCGCACTTTGGCTTTGGCCGAGAAATCAACCCAGCAGAGTGA 360	841 GCACCTCTGACAGGACACTGGTGGCTCTCCAGTTCTCTCTCTCTGACAAAGAGCT 900
DB	3957 GGAAGACATTAATAATTCGCACTTTGGCTTTGGCCGAGAAATCAACCCAGCAGAGTGA 3916	4397 GCACCTCTGACAGGACACTGGTGGCTCTCCAGTTCTCTCTCTCTGACAAAGAGCT 4456
QY	361 GTTCAGCAGTACGGCTCCCTGAGTTCTGCTCTCCCGAGATCATCCAGCAGAACCTCTG 420	901 CGCCCATTTGCCCCGGCTAAAGTCACTGCACCTCCCGGTGACACACTCACCATGCT 960
DB	3917 GTTCAGCAGTACGGCTCCCTGAGTTCTGCTCTCCCGAGATCATCCAGCAGAACCTCTG 3976	4457 CGCCCATTTGCCCCGGCTAAAGTCACTGCACCTCCCGGTGACACACTCACCATGCT 4516
QY	421 GAGCGAAGCCTCGACATTTGGGCCATGGGTGTCACTCTACCTCAGCCTGACCTGTCTC 480	961 GCACCCCGGGGGTCTCTGGGCCCTTCGGCCAGCTCTGTGAGGAAGCCGAGGCCAGTGA 1020
DB	3977 GAGCGAAGCCTCGACATTTGGGCCATGGGTGTCACTCTACCTCAGCCTGACCTGTCTC 4036	4517 GCACCCCGGGGGTCTCTGGGCCCTTCGGCCAGCTCTGTGAGGAAGCCGAGGCCAGTGA 4576
QY	481 ATCCCATTTGCGCGGAGAGTGACCGTGCACCCCTCCTGAGAGTCTCTGGAGGGGGCGT 540	1021 GCGCTCCACCGAGGGCCCAAGCTCGCTGCATCTCCCGAGGGTGCCTGGGCCACCCGCCG 1080
DB	4037 ATCCCATTTGCGCGGAGAGTGACCGTGCACCCCTCCTGAGAGTCTCTGGAGGGGGCGT 4096	4577 GCGCTCCACCGAGGGCCCAAGCTCGCTGCATCTCCCGAGGGTGCCTGGGCCACCCGCCG 4636

Db 5177 TTCTGCCCCCAGGAGGCTGCAGCCCCCACCAGCAGTTGCCCATGCCCTCTCTGGCTC 5236
Qy 1681 CTTCCCTCCAGGATCTTCAAGAGGCCCCCTTAGTACCTCAAGCCCCCTTCTTGGGACA 1740
Db 5237 CTTCCCTCCAGGATCTTCAAGAGGCCCCCTTAGTACCTCAAGCCCCCTTCTTGGGACA 5296
Qy 1741 GCCCAGGACCCCTGCCCCCTGCACAAAGAGCCGCCCATTTGACTCTAAGATGGGGCC 1800
Db 5297 GCCCAGGACCCCTGCCCCCTGCACAAAGAGCCGCCCATTTGACTCTAAGATGGGGCC 5356
Qy 1801 TGGAGACATCTCTTCTCTGGAGGCCCCAAACCCGGCCCCCTGCACTGCCAGGGTCAGC 1860
Db 5357 TGGAGACATCTCTTCTCTGGAGGCCCCAAACCCGGCCCCCTGCACTGCCAGGGTCAGC 5416
Qy 1861 CTCACGCGAGCTCTTCCCAAGTAGCTTCCCTCAGGGTGGGCTCTCCCAAGGTGGGCAC 1920
Db 5417 CTCACGCGAGCTCTTCCCAAGTAGCTTCCCTCAGGGTGGGCTCTCCCAAGGTGGGCAC 5476
Qy 1921 AAGGCTGGCCCCCTCTGAGTGGGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 1980
Db 5477 AAGGCTGGCCCCCTCTGAGTGGGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 5536
Qy 1981 CTCACACCCACCTTGACGCGGCTCAGCAACAGGCGACCATGCGCAAGTTCTCCCTGGG 2040
Db 5537 CTCACACCCACCTTGACGCGGCTCAGCAACAGGCGACCATGCGCAAGTTCTCCCTGGG 5596
Qy 2041 TGGTCGGGGGGCTACGACGCGTGGCTGGCTATGGCACTTTGCTTTGGTGGAGATGC 2100
Db 5597 TGGTCGGGGGGCTACGACGCGTGGCTGGCTATGGCACTTTGCTTTGGTGGAGATGC 5656
Qy 2101 AGGGGGCATGTGGGGCAGGGGCCCATGTGGCCAGGATAGCTGGGCTGTGTCCAGTC 2160
Db 5657 AGGGGGCATGTGGGGCAGGGGCCCATGTGGCCAGGATAGCTGGGCTGTGTCCAGTC 5716
Qy 2161 GGAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Db 5717 GGAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5776
Qy 2221 CAGGCTGAGAGGCCACTGCCCCAGGTCAGTCAGAGGCTGTGCTCAGAGGTCGGCAGGGC 2280
Db 5777 CAGGCTGAGAGGCCACTGCCCCAGGTCAGTCAGAGGCTGTGCTCAGAGGTCGGCAGGGC 5836
Qy 2281 TCCACACGAGGCTCTCAGAGGCCACCCCATGGAGGACATCGGGCAGGTCCTCCCTGT 2340
Db 5837 TCCACACGAGGCTCTCAGAGGCCACCCCATGGAGGACATCGGGCAGGTCCTCCCTGT 5896
Qy 2341 GCAGATCCGGACCTGTTCAGGTGATGCGAGGCGCGGCGACACATATCCCTGGACATTC 2400
Db 5897 GCAGATCCGGACCTGTTCAGGTGATGCGAGGCGCGGCGGCGACACATATCCCTGGACATTC 5956
Qy 2401 CGAGGTGGACCCCGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
Db 5957 CGAGGTGGACCCCGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 6016
Qy 2461 CGAGTTTATGATCTTCAGAAAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCCCAT 2520
Db 6017 CGAGTTTATGATCTTCAGAAAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCCCAT 6076
Qy 2521 GGCTGAGGAGGCTGGCGGAGTTCCCGAGCCACAGTGGCCCTGGCCAGGTAAGTGGG 2580
Db 6077 GGCTGAGGAGGCTGGCGGAGTTCCCGAGCCACAGTGGCCCTGGCCAGGTAAGTGGG 6136
Qy 2581 CCCCCACGAGGCTGGAGATCACAGAGGAGTCAGAGGATGGACGCGCTGTGGCAGA 2640
Db 6137 CCCCCACGAGGCTGGAGATCACAGAGGAGTCAGAGGATGGACGCGCTGTGGCAGA 6196
Qy 2641 GGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2700
Db 6197 GGCTGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 6256
Qy 2701 TGGGAGGACCTGGCGCTGGATGAGCTGCAGAGCTGGGGCTGGCTGAGAGGTGAAGGC 2760

Db 6257 TGGGAGGACCTGGCGCTGGATGAGCCCTGCAGAGCTGGGGCTGGAGAGTGAAGGC 6316
Qy 2761 CTCGTGTGAGCACATCTCCCGATCTCGAAGAGGCGAGCGGAAAGTCTTGGAGAGGAGG 2820
Db 6317 CTCGTGTGAGCACATCTCCCGATCTCGAAGAGGCGAGCGGAAAGTCTTGGAGAGGAGG 6376
Qy 2821 GCCCCCGAGGAAGAGCCAGGCTTGTCTTCCGCTCTCAGGCTCTGAAGAGCTGGGA 2880
Db 6377 GCCCCCGAGGAAGAGCCAGGCTTGTCTTCCGCTCTCAGGCTCTGAAGAGCTGGGA 6436
Qy 2881 CCGAGCGCGACATCTTCTTAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCTCAGT 2940
Db 6437 CCGAGCGCGACATCTTCTTAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCTCAGT 6496
Qy 2941 GACACTGGCTTGCAGAGTGTGAGCCAGCCAGCTGCCAGCCACCTCGAGCAAGACGG 3000
Db 6497 GACACTGGCTTGCAGAGTGTGAGCCAGCCAGCTGCCAGCCACCTCGAGCAAGACGG 6556
Qy 3001 A 3001
Db 6557 A 6557

RESULT 3
ABX11642
ID ABX11642 standard; cDNA; 24120 BP.
XX
AC ABX11642;
XX
DT 09-MAY-2003 (first entry)
XX
DE Human serine/threonine or protein kinase 12599, cDNA.
XX
KW Human; ss; gene; serine/threonine kinase; protein kinase; 12599;
cardiovascular disease; heart failure; myocardial infarction;
blood vessel disorder; atherosclerosis; Kaposi's sarcoma;
blood platelet disorder; thrombocytopenia; leukaemia; Hodgkin's disease;
haemolytic anaemia; cellular proliferative disorder; cancer;
protein kinase disorder; autoimmune disorder; diabetes mellitus;
psoriasis; inflammatory bowel disease; rheumatoid arthritis;
multiple sclerosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..71
FT CDS /*tag= a
72..23978
FT /*tag= b
FT /product= "Kinase 12599"
FT /note= "This CDS is specifically claimed in claim 2"
FT 3'UTR 23979..24120
FT /*tag= c
XX
PN US2002168742-A1.
XX
PD 14-NOV-2002.
XX
PF 15-FEB-2002; 2002US-00077130.
XX
PR 15-FEB-2001; 2001US-0269201P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Kapeller-Libermann R, Acton SL;
XX
DR WPI; 2003-298729/29.
XX
PT P-PSDB; ABG76187.
XX
PT Novel isolated human protein kinase, designated 59079 or 12599
polypeptide, useful as diagnostic and therapeutic agents for preventing
cardiovascular diseases, proliferative disorders, and protein kinase
disorders.

XX Claim 2; Page 58-84; 119pp; English.

XX The invention relates to an isolated human serine/threonine or protein

CC kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule

CC comprising at least 85% identity to the nucleic acids appearing as

CC ABX11641 and ABX11642 or their complement, a naturally occurring variant

CC of the kinases or their fragments. Also included are a non-human host

CC cell containing the nucleic acids, an antibody specific for the proteins,

CC identifying a compound which binds to the kinase (by contacting the

CC kinase or a cell expressing the kinase with a test compound and

CC determining whether the kinase binds to the test compound) and modulating

CC the activity of kinase using the identified compound. The kinases and

CC their encoding nucleic acids are useful as diagnostic and therapeutic

CC agents for preventing a disease or condition associated with an aberrant

CC or unwanted 59079 or 12599 activity in a subject, including

CC cardiovascular diseases such as heart failure, and myocardial infarction;

CC disorders involving blood vessels such as atherosclerosis, and Kaposi's

CC sarcoma; blood platelets disorder such as thrombocytopenia, leukaemia,

CC Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders

CC such as cancer; and protein kinase disorders such as autoimmune

CC disorders, diabetes mellitus, psoriasis, inflammatory bowel disease,

CC rheumatoid arthritis, and multiple sclerosis (many examples of diseases

CC and disorders are included in the specification). The kinases, their

CC encoding nucleic acids and antibodies are useful in screening assays,

CC detection assays (e.g. forensic biology), and predictive medicine (e.g.

CC diagnostic assays, prognostic assays, and monitoring clinical trials and

CC pharmacogenomics). The kinases and their encoding nucleic acids are

CC useful as query sequences to perform a search against public databases to

CC identify other family members or related sequences. The present sequence

CC encodes the kinase 12599

XX

SQ Sequence 24120 BP; 4719 A; 7265 C; 8278 G; 3858 T; 0 U; 0 Other;

Query Match 100.0%; Score 3001; DB 8; Length 24120;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CCTACGGAGCAGAACTCGGGCCAGGAGATACAGGGAGCGAGACATCTCGCGCGCTGAG 60

Db 19571 CCTACGGAGCAGAACTCGGGCCAGGAGATACAGGGAGCGAGACATCTCGCGCGCTGAG 19630

Qy 61 CCACCCGCTGTCACGGGGCTGACGACCTGACCTGACCTGACCTGACCTGACCTGACCTGAC 120

Db 19631 CCACCCGCTGTCACGGGGCTGACGACCTGACCTGACCTGACCTGACCTGACCTGACCT 19690

Qy 121 CCTGGAGCTGTCTCATCCGAGGAGCTGCTGGACCGCTGTACAGGAAGGCGGTGGTAC 180

Db 19691 CCTGGAGCTGTCTCATCCGAGGAGCTGCTGGACCGCTGTACAGGAAGGCGGTGGTAC 19750

Qy 181 GGAGGCGAGGTCAAGGTCTACATCCAGAGCTGCTGGAGGGGCTGACCTGACACAG 240

Db 19751 GGAGGCGAGGTCAAGGTCTACATCCAGAGCTGCTGGAGGGGCTGACCTGACACAG 19810

Qy 241 CCATGGGCTTCTCCACCTGGAGCATAAAGCCCTCTAATCTGATGGTGTGATCTGCGCG 300

Db 19811 CCATGGGCTTCTCCACCTGGAGCATAAAGCCCTCTAATCTGATGGTGTGATCTGCGCG 19870

Qy 301 GGAAGACATTAATAATCTGCGACTTTGGCTTTGGCCAGAAATCAACCCAGCAGAGCTGCA 360

Db 19871 GGAAGACATTAATAATCTGCGACTTTGGCTTTGGCCAGAAATCAACCCAGCAGAGCTGCA 19930

Qy 361 GTTCAGGACGATCGGCTCCCTGAGTTGCTGCTCCCGGAGATCATCCAGCAGAACCTGT 420

Db 19931 GTTCAGGACGATCGGCTCCCTGAGTTGCTGCTCCCGGAGATCATCCAGCAGAACCTGT 19990

Qy 421 GAGCAAGCCTCCGACATTTGGGCGATGGGTGTCTCTCTACCTCAGCCTGACCTGCTC 480

Db 19991 GAGCAAGCCTCCGACATTTGGGCGATGGGTGTCTCTCTACCTCAGCCTGACCTGCTC 20050

Qy 481 ATCCCAATTTGCGGCGAGAGTGAACCGTGCACCCCTCTCTGAAAGCTCTTGGAGGGGCGGT 540

Db 20051 ATCCCAATTTGCGGCGAGAGTGAACCGTGCACCCCTCTCTGAAAGCTCTTGGAGGGGCGGT 20110

Qy 541 GTCATGAGAGAGCCCATGGCTGCCACCTCAGCGAAGAGCGCCAAAGACTTTCATCAAGGC 600

Db 20111 GTCATGAGAGAGCCCATGGCTGCCACCTCAGCGAAGAGCGCCAAAGACTTTCATCAAGGC 20170

Qy 601 TAGCCTCGAGAGAGCCCTCAGGCCCGGCTAGTGGCGGCCAGTGCCTCTCCCAACCCCTG 660

Db 20171 TAGCCTCGAGAGAGCCCTCAGGCCCGGCTAGTGGCGGCCAGTGCCTCTCCCAACCCCTG 20230

Qy 661 GTTCTCAGATCCATGCTCGGAGGAGGCGCCACTTTCATCAACCAAGCAGCTCAAGTT 720

Db 20231 GTTCTCAGATCCATGCTCGGAGGAGGCGCCACTTTCATCAACCAAGCAGCTCAAGTT 20290

Qy 721 CCTCTCGCCGGAAGTCCCTGGGAGCGCTCCCTGATGAGCTTACAAGTCCATCTCTGGTGAT 780

Db 20291 CCTCTCGCCGGAAGTCCCTGGGAGCGTTTCCCTGATGAGTCAAGTTCATCTCTGGTGAT 20350

Qy 781 GCGCTCCATCTCTGAGTGTGCGGGGCCACCCGAGCAGCCCTCTCTCGCGGTAGCCCG 840

Db 20351 GCGCTCCATCTCTGAGTGTGCGGGGCCACCCGAGCAGCCCTCTCTCGCGGTAGCCCG 20410

Qy 841 GCACCTCTGAGGAGCACTGGTGGCTCTCCAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900

Db 20411 GCACCTCTGAGGAGCACTGGTGGCTCTCTCCAGTTCTCTCTCTCTCTCTCTCTCTCTCTCT 20470

Qy 901 CGCCCTCATTTGCGCGGCTAAGTCACTGCCACCTCTCCCGGTGACACACTCACCACCTGCT 960

Db 20471 CGCCCTCATTTGCGCGGCTAAGTCACTGCCACCTCTCCCGGTGACACACTCACCACCTGCT 20530

Qy 961 GCACCCCGGGGCTTCTGCGGCCCTGCGGCGCTGCGGAGCCGAGGAGCCGAGGAGTGA 1020

Db 20531 GCACCCCGGGGCTTCTGCGGCCCTGCGGCGCTGCGGAGCCGAGGAGCCGAGGAGTGA 20590

Qy 1021 GCCTCCACCGAGGCCCGGAGCTCGGCTCGGCTCATCTCCGAGGCTGCGGGGCGCCGCGCG 1080

Db 20591 GCCTCCACCGAGGCCCGGAGCTCGGCTCGGCTCATCTCCGAGGCTGCGGGGCGCCGCGCGCG 20650

Qy 1081 CAGGGCTGCTGCTCCCGGCGACAGCGCTCATCCGAGCGCTGTTCTACACAGCGGGGTGA 1140

Db 20651 CAGGGCTGCTGCTCCCGGCGACAGCGCTCATCCGAGCGCTGTTCTACACAGCGGGGTGA 20710

Qy 1141 GAGCCCTGAGCAGCGGGCCCTGCGCCCGGGGAGCAGCGGCGACCCCGGCGCGCGGCGCA 1200

Db 20711 GAGCCCTGAGCAGCGGGCCCTGCGCCCGGGGAGCAGCGGCGACCCCGGCGCGCGGCGCA 20770

Qy 1201 CTTGCTGAAGGGCGGCTACATTGCGGGGCGCTGCCAGCGCTCGCGGAGCCTGATGGA 1260

Db 20771 CTTGCTGAAGGGCGGCTACATTGCGGGGCGCTGCCAGCGCTCGCGGAGCCTGATGGA 20830

Qy 1261 GCACCCGCTGCTGGAGGAGGCGCCCGAGGAGGAGCGGCGACCCCTCTCTGCGCCAAAGC 1320

Db 20831 GCACCCGCTGCTGGAGGAGGCGCCCGAGGAGGAGCGGCGACCCCTCTCTGCGCCAAAGC 20890

Qy 1321 CCCTCATTTGAGAGTCCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1380

Db 20891 CCCTCATTTGAGAGTCCCTCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 20950

Qy 1381 CAGCCACTCCCTGGAAACATGACTCTCCGAGCAGCCCGCGCCCTCTCTCGAGGCGCTGCGG 1440

Db 20951 CAGCCACTCCCTGGAAACATGACTCTCCGAGCAGCCCGCGCCCTCTCTCGAGGCGCTGCGG 21010

Qy 1441 TGAGGACAGCAGCTGCTTTCAGCCCTCTCGGGGGGGGCGCCCTATCAGGGACATGCGGGCA 1500

Db 21011 TGAGGACAGCAGCTGCTTTCAGCCCTCTCGGGGGGGGCGCCCTATCAGGGACATGCGGGCA 21070

Qy 1501 CCCTCAGGGCTCCAGCAGCTTCCATCCACTGGTGGCGACCCGAGCAGCTGCTCAGCCAG 1560

Db 21071 CCCTCAGGGCTCCAGCAGCTTCCATCCACTGGTGGCGACCCGAGCAGCTGCTCAGCCAG 21130

Qy 1561 GAGGCCATCCCGGAGCAGCCCTTGGGGGCGAGCCCGCTTCTGCGACCCCGAAGCAGGG 1620

Db 21131 GAGGCCATCCCGGAGCAGCCCTTGGGGGCGAGCCCGCTTCTGCGACCCCGAAGCAGGG 21190

Qy 1621 TTCTGCCCCCAGAGGGCTGCAGCCCCCAGAGAGTTGCCCATGCCCCCTCTGGCTC 1680
Db 21191 TTCTGCCCCCAGAGGGCTGCAGCCCCCAGAGAGTTGCCCATGCCCCCTCTGGCTC 21250
Qy 1681 CTTCCCTCCAGGATCTTTGAAAGAGGGCCCCCTTAGTACCCCTCAAGAGCCCCCTTTTGGGACA 1740
Db 21251 CTTCCCTCCAGGATCTTTGAAAGAGGGCCCCCTTAGTACCCCTCAAGAGCCCCCTTTTGGGACA 21310
Qy 1741 GCCCCAGGACCCCTGCTGCTGCAAGAGAGCCCCCATTTGAGACTTAAGATGGGGCC 1800
Db 21311 GCCCCAGGACCCCTGCTGCTGCAAGAGAGCCCCCATTTGAGACTTAAGATGGGGCC 21370
Qy 1801 TGGAGATCTCTCTCTCTGGAGGCCAAGACCCGGCCCCCTGCAGTTCCCCAGGGTCAGC 1860
Db 21371 TGGAGATCTCTCTCTCTGGAGGCCAAGACCCGGCCCCCTGCAGTTCCCCAGGGTCAGC 21430
Qy 1861 CTCCAGGCGAGCTCTTCCCAAGTAGCTCTCCTCAGGGTGGGCTCCTCCAGAGTGGGCAC 1920
Db 21431 CTCCAGGCGAGCTCTTCCCAAGTAGCTCTCCTCAGGGTGGGCTCCTCCAGAGTGGGCAC 21490
Qy 1921 AGAGCTTGGCCCCCTCTGAGATGGAGGGCTGGAAGAGAGGCTGAGGATCTGTCCGA 1980
Db 21491 AGAGCTTGGCCCCCTCTGAGATGGAGGGCTGGAAGAGAGGCTGAGGATCTGTCCGA 21550
Qy 1981 CTCCACACCCACCTTGAGCGGCTCTCAGAACAGGCGACCATGCGAAGTTCTCCCTGGG 2040
Db 21551 CTCCACACCCACCTTGAGCGGCTCTCAGAACAGGCGACCATGCGAAGTTCTCCCTGGG 21610
Qy 2041 TGGTGGGGGGCTACGAGGCGTGGCTGGCTATGGCACTTTGGTGGAGATGC 2100
Db 21611 TGGTGGGGGGCTACGAGGCGTGGCTGGCTATGGCACTTTGGTGGAGATGC 21670
Qy 2101 AGGGGGGATGCTGGGGCAGGGGCCATGTGGGCAGAGATAGCCTTGGGCTGTGTCCAGTC 2160
Db 21671 AGGGGGGATGCTGGGGCAGGGGCCATGTGGGCAGAGATAGCCTTGGGCTGTGTCCAGTC 21730
Qy 2161 GGAGGAGGAGCAGAGGAGGCGCAGGGCTGAGTCCAGTCSGAGGAGCAGGAGGC 2220
Db 21731 GGAGGAGGAGCAGAGGAGGCGCAGGGCTGAGTCCAGTCSGAGGAGCAGGAGGC 21790
Qy 2221 CAGGGCTGAGAGCCACTGCCCCAGGTCAAGTCAAGGCGCTGTGCTGAGGTGCGCAGGGC 2280
Db 21791 CAGGGCTGAGAGCCACTGCCCCAGGTCAAGTCAAGGCGCTGTGCTGAGGTGCGCAGGGC 21850
Qy 2281 TCCACACGAGAGCTCTCCAGAGCCACCCCATGAGGAGACATCGGGCAGGTCTCCCTGGT 2340
Db 21851 TCCACACGAGAGCTCTCCAGAGCCACCCCATGAGGAGACATCGGGCAGGTCTCCCTGGT 21910
Qy 2341 GCAGATCGGAGCCTGTCAAGTGTGCGGAGCGCGCGACACATATCCCTGGACATTC 2400
Db 21911 GCAGATCGGAGCCTGTCAAGTGTGCGGAGCGCGCGACACATATCCCTGGACATTC 21970
Qy 2401 CGAGGTGAGACCCCGCTACCTCAACCTCTCAGACTGTACGATATCAAGTACCTCCCAT 2460
Db 21971 CGAGGTGAGACCCCGCTACCTCAACCTCTCAGACTGTACGATATCAAGTACCTCCCAT 22030
Qy 2461 CGAGTTTATGATTTTCAAGAAAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCCCAT 2520
Db 22031 CGAGTTTATGATTTTCAAGAAAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCCCAT 22090
Qy 2521 GGCTGAGGAGAGCTGGCGAGTTTCCCGAGCCACAGTGGCCCTGGCAGGTGAATGGG 2580
Db 22091 GGCTGAGGAGAGCTGGCGAGTTTCCCGAGCCACAGTGGCCCTGGCAGGTGAATGGG 22150
Qy 2581 CCCCCAGCGGCTGGAGATCACAGAGGAGTCAAGAGATGTGGACGGCTGTCTGGCAGA 2640
Db 22151 CCCCCAGCGGCTGGAGATCACAGAGGAGTCAAGAGATGTGGACGGCTGTCTGGCAGA 22210
Qy 2641 GGCTGCCGTGGCAGGAGCGCAAGTGTCTCTCGCGTCAAGAGCTCTTCCACTTCCC 2700
Db 22211 GGCTGCCGTGGCAGGAGCGCAAGTGTCTCTCGCGTCAAGAGCTCTTCCACTTCCC 22270
Qy 2701 TGGAGGACCTGCGGCTGGATGAGCCTGCAGAGCTGGGGCTGGTGAGAGAGTGAAGGC 2760

Db 22271 TGGAGGACCTGCGGCTGGATGAGCCTGCAGAGCTGGGGCTGGTGAGAGTGAAGGC 22330
Qy 2761 CTCCGTGGAGCACATCTCCCGGATCTTAAGGGCAGGCCGGAAGGTCTGGAGAAGGAGG 2820
Db 22331 CTCCGTGGAGCACATCTCCCGGATCTTAAGGGCAGGCCGGAAGGTCTGGAGAAGGAGG 22390
Qy 2821 GCCCCCAGGAGAGAGCAGGCTTGTCTTCCGCTCTCAGAGTCTGAAGAGCTGGGA 2880
Db 22391 GCCCCCAGGAGAGAGCAGGCTTGTCTTCCGCTCTCAGAGTCTGAAGAGCTGGGA 22450
Qy 2881 CCAGAGCCCGACATCTCTTAAGGGAGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 2940
Db 22451 CCAGAGCCCGACATCTCTTAAGGGAGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 22510
Qy 2941 GACACTGCGCTGCCAGGTGTGAGCCAGCCAGCTGCCAGGCCACCTGGAGCAAGACGG 3000
Db 22511 GACACTGCGCTGCCAGGTGTGAGCCAGCCAGCTGCCAGGCCACCTGGAGCAAGACGG 22570
Qy 3001 A 3001
Db 22571 A 22571

RESULT 4

AAC62286

ID AAC62286 standard; cDNA; 5007 BP.

AC AAC62286;

XX

AC AAC62286;

XX

DT 19-MAR-2001 (first entry)

XX

XX

DE cDNA encoding a full length human signal transduction polypeptide.

XX

XX

KW Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;

KW congestive heart failure; dilated congestive cardiomyopathy;

KW hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;

KW mitral valve disease; aortic valve disease; tricuspid valve disease;

KW myocardial infarction; cardiac arrhythmia; arteriosclerosis;

KW atherosclerosis; cardiac tumour; microbial infection; ss.

XX

OS Homo sapiens.

XX

FH Key

CDS

Location/Qualifiers

97..4929

/*tag= a

/product= "signal transduction polypeptide H19G5"

XX

FN WO200063381-A1.

XX

XX

PD 26-OCT-2000.

XX

XX

PF 11-APR-2000; 2000WO-US009488.

XX

XX

PR 16-APR-1999; 99US-0129553P.

XX

XX

(SCIO-) SCIOS INC.

PA

Zeng W, Stanton L, Kong H;

XX

XX

DR WPI; 2001-007013/01.

XX

DR P-PSDB; AAB30568.

XX

XX

PT Novel h19G5 polypeptides capable of regulating signal transduction and

PT exhibiting kinase activity useful for identifying antibodies to treat

PT cardiac diseases, and additional mediators of signal transduction.

XX

XX

PS Claim 4; Page 59-61; 81pp; English.

XX

XX

CC The present sequence encodes a human protein with putative function in

CC signal transduction. The polypeptide is designated H19G5. The protein is

CC capable of regulating signal transduction and exhibits kinase activity.

CC The H19G5 transcript is expressed in the heart. H19G5 polypeptides and

polynucleotides are useful for preventing or treating a cardiac disease, such as congestive heart failure, dilated congestive cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy, mitral valve disease, aortic valve disease or tricuspid valve disease, angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary or rheovascular hypertension, arteriosclerosis, atherosclerosis and cardiac tumours in humans. The polypeptide is also useful for detecting the expression of a protein capable of regulating signal transduction or the expression of a protein capable of acting as a donor or acceptor molecule of a phosphate group. The monoclonal antibodies can be used as probes for detecting discrete antigens expressed by tissue or cell samples, and therefore used in humans for localization and monitoring of microbial infection

SQ Sequence 5007 BP; 946 A; 1722 C; 1541 G; 798 T; 0 U; 0 Other;

Query Match	99.9%;	Score 2999.4;	DB 5;	Length 5007;	
Best Local Similarity	100.0%;	Prod. No. 0;			
Matches 3000;	Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	CCTACGGAGCAAACTCGGGCCCAAGGCATACAGGAGCGAGACATCTGGCCGGCGCTGAG	60		
Db	522	CCTACGGAGCAAACTCGGGCCCAAGGCATACAGGAGCGAGACATCTGGCCGGCGCTGAG	581		
Qy	61	CCACCCGCTGGTCACGGGGCTGCTGGACAGCTTTGAGACCCGCAAGACCCCTCATCCTCAT	120		
Db	582	CCACCCGCTGGTCACGGGGCTGCTGGACAGCTTTGAGACCCGCAAGACCCCTCATCCTCAT	641		
Qy	121	CCTGGAGCTGTGCTTCATCCGAGGAGCTGCTGGACCGCCTGTATACGGAAGGCGCTGGTGA	180		
Db	642	CCTGGAGCTGTGCTTCATCCGAGGAGCTGCTGGACCGCCTGTATACGGAAGGCGCTGGTGA	701		
Qy	181	GGAGGCCGAGGCTCAAAGTCTTACATCCAGCAGCTGGTGGAGGGGCTGCACCTACCTGCACAG	240		
Db	702	GGAGGCCGAGGCTCAAAGTCTTACATCCAGCAGCTGGTGGAGGGGCTGCACCTACCTGCACAG	761		
Qy	241	CCATGGCGTTCTCCACTCGGACATAAAGCCCTTAAACATCCTGATGTTGATCCTTGCCCG	300		
Db	762	CCATGGCGTTCTCCACTCGGACATAAAGCCCTTAAACATCCTGATGTTGATCCTTGCCCG	821		
Qy	301	GGAAAGCATTTAAATCTGGCACTTTGGCTTTGGCCAGAAACATCACCCACAGAGCTGCA	360		
Db	822	GGAAAGCATTTAAATCTGGCACTTTGGCTTTGGCCAGAAACATCACCCACAGAGCTGCA	881		
Qy	361	GTTCAGCGAGTACGGCTCCCTCAGTTCGTCTCCCCGAGATCATCAGCAGAAACCTGT	420		
Db	882	GTTCAGCGAGTACGGCTCCCTCAGTTCGTCTCCCCGAGATCATCAGCAGAAACCTGT	941		
Qy	421	GACGGAAGCCTCCGACATTTGGCCATGGGTGTTCATCTCCTACCTCAGCCTGACCTGTCTC	480		
Db	942	GACGGAAGCCTCCGACATTTGGCCATGGGTGTTCATCTCCTACCTCAGCCTGACCTGTCTC	1001		
Qy	481	ATCCCATTTGCGGGGAGTAGTACCGTGCACCCCTCCTGAACGTCCTGGAGGGGGCGCT	540		
Db	1002	ATCCCATTTGCGGGGAGTAGTACCGTGCACCCCTCCTGAACGTCCTGGAGGGGGCGCT	1061		
Qy	541	GTCATGGAGCGCCCATGGCTGCCACCTCAGCAGGAAGCGCCAAAGATTTCATCAAGGC	600		
Db	1062	GTCATGGAGCGCCCATGGCTGCCACCTCAGCAGGAAGCGCCAAAGATTTCATCAAGGC	1121		
Qy	601	TACGCTGCAGAGAGCCCTCAGGCCCGGCTAGTGGGGCCAGTGGCTCTCCACACCCCTG	660		
Db	1122	TACGCTGCAGAGAGCCCTCAGGCCCGGCTAGTGGGGCCAGTGGCTCTCCACACCCCTG	1181		
Qy	661	GTTCTCTGAATTCATGCTTCGGAGGAGGCCCATTCATCAACACCAAGCAGCTCAAGTT	720		
Db	1182	GTTCTCTGAATTCATGCTTCGGAGGAGGCCCATTCATCAACACCAAGCAGCTCAAGTT	1241		
Qy	721	CCTCTCTGGCCGAAGTCGCTGGCAGGTTCCCTGTATGAGCTACAGTCCATCCTGGTGTAT	780		
Db	1242	CCTCTCTGGCCGAAGTCGCTGGCAGGTTCCCTGTATGAGCTACAGTCCATCCTGGTGTAT	1301		
Qy	781	GGGCTCCATTCCTGTAGTGTCTCGGGGGGCCACCCGACAGCCCTCCCTCGGCGTAGTGC	840		

Db 2382 CTCCAGGCGAGCTCTTCCCAAGTGAGCTCCCTCAGGGTGGGCTCCTCCAGGTGGGCAC 2441
Qy 1921 AGAGCTTGGCCCTCCCTGATGCGAGGCTGACCCAGAGGCTGAGATCTGTCCGA 1980
Db 2442 AGAGCTTGGCCCTCCCTGATGCGAGGCTGACCCAGAGGCTGAGATCTGTCCGA 2501
Qy 1981 CTCCACACCCACCTTGAGCGGCTCAGGACAGCGGACCATGCGAGGTTCTCCCTGGG 2040
Db 2502 CTCCACACCCACCTTGAGCGGCTCAGGACAGCGGACCATGCGAGGTTCTCCCTGGG 2561
Qy 2041 TGGTCTGGGGGGCTACGCGAGCGTGGCTGAGTCCGACACCTTTGCTTTGGTGAGATGC 2100
Db 2562 TGGTCTGGGGGGCTACGCGAGCGTGGCTGAGTCCGACACCTTTGCTTTGGTGAGATGC 2621
Qy 2101 AGGGGGATGCTGGGGGCGGCGCATGTGGGCGAGATAGCTGGGCTGTGTCCAGTC 2160
Db 2622 AGGGGGATGCTGGGGGCGGCGCATGTGGGCGAGATAGCTGGGCTGTGTCCAGTC 2681
Qy 2161 GGAGGAGGAGCAGGAGGAGGCGCAGGCTGAGTCCGAGTGGAGGAGCAGCAGGAGGC 2220
Db 2682 AGAGGAGGAGCAGGAGGAGGCGCAGGCTGAGTCCGAGTGGAGGAGCAGCAGGAGGC 2741
Qy 2221 CAGGGCTGAGAGCCACTGCCCCAGGTCAGTGCAGGCGCTGTGCTGAGGTGGCAGGCG 2280
Db 2742 CAGGGCTGAGAGCCACTGCCCCAGGTCAGTGCAGGCGCTGTGCTGAGGTGGCAGGCG 2801
Qy 2281 TCCACACGAGGCTCTCCAGAGCCACCCATGGAGGACATCGGGCAGGTCTCCCTGGT 2340
Db 2802 TCCACACGAGGCTCTCCAGAGCCACCCATGGAGGACATCGGGCAGGTCTCCCTGGT 2861
Qy 2341 GCAGATCCGGGACCTGTGAGTGTGCGGAGCGGCGGACACATATCCCTGGACATTTTC 2400
Db 2862 GCAGATCCGGGACCTGTGAGTGTGCGGAGCGGCGGACACATATCCCTGGACATTTTC 2921
Qy 2401 CGAGGTGACCCCGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTT 2460
Db 2922 CGAGGTGACCCCGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTT 2981
Qy 2461 CGAGTTTATGATCTTCAAGAAAGTCCCAAGTCCGCTCAGCAGCGCCCTCCCCCAT 2520
Db 2982 CGAGTTTATGATCTTCAAGAAAGTCCCAAGTCCGCTCAGCAGCGCCCTCCCCCAT 3041
Qy 2521 GGCTGAGGAGAGCTGGCGAGTTCCCGAGCCACGTCGGCCCTGGCCAGGTGAATGGG 2580
Db 3042 GGCTGAGGAGAGCTGGCGAGTTCCCGAGCCACGTCGGCCCTGGCCAGGTGAATGGG 3101
Qy 2581 CCCCACGCGGCTGGAGATCACAGAGGAGTCAGAGGATGTGACGCGCTGTGGCAGA 2640
Db 3102 CCCCACGCGGCTGGAGATCACAGAGGAGTCAGAGGATGTGACGCGCTGTGGCAGA 3161
Qy 2641 GGCTGCGGTGGCAGGAGCGCAAGTGGTCTCGCCGCTCAGCGAGCTCTTTCCACTTCCC 2700
Db 3162 GGCTGCGGTGGCAGGAGCGCAAGTGGTCTCGCCGCTCAGCGAGCTCTTTCCACTTCCC 3221
Qy 2701 TGGAGGACCTGGCGCTGGATGAGCTGCGAGAGCTGGGGCTGGGTGAGAGGTGAAGGC 2760
Db 3222 TGGAGGACCTGGCGCTGGATGAGCTGCGAGAGCTGGGGCTGGGTGAGAGGTGAAGGC 3281
Qy 2761 CTCCGTGGAGCACATCTCCGGATCTTGAAGGCGAGGCGGAGGTCTGGAGAGGAGGG 2820
Db 3282 CTCCGTGGAGCACATCTCCGGATCTTGAAGGCGAGGCGGAGGTCTGGAGAGGAGGG 3341
Qy 2821 GCCCCCAGGAGAGCCAGGCTTGTCTTCCCTCCGCTCTCAGGTCTGAAGAGCTGGGA 2880
Db 3342 GCCCCCAGGAGAGCCAGGCTTGTCTTCCCTCCGCTCTCAGGTCTGAAGAGCTGGGA 3401
Qy 2881 CCGAGGCGGACATCTCAAGGAGCTCTCAGATGAGACTGTGGTCTGGGCCAGTCAGT 2940
Db 3402 CCGAGGCGGACATCTCAAGGAGCTCTCAGATGAGACTGTGGTCTGGGCCAGTCAGT 3461
Qy 2941 GACACTGGCTGCGAGGTGTGAGCCCGAGCTGCCCCAGGCGCAGCTGGAGCAAGAGCGG 3000
Db 3462 GACACTGGCTGCGAGGTGTGAGCCCGAGCTGCCCCAGGCGCAGCTGGAGCAAGAGCGG 3521

Qy 3001 A 3001
Db 3522 A 3522

RESULT 5

AAC62287
ID AAC62287 standard; cDNA; 7928 BP.

XX AAC62287;

AC AC
DT 19-MAR-2001 (first entry)

XX cDNA encoding a splice variant of a signal transduction polypeptide.

DE Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
XX congestive heart failure; dilated congestive cardiomyopathy;
KW hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
KW mitral valve disease; aortic valve disease; tricuspid valve disease;
KW myocardial infarction; cardiac arrhythmia; arteriosclerosis;
KW atherosclerosis; cardiac tumour; microbial infection; splice variant; ss.

OS Homo sapiens.

XX Location/Qualifiers

Key 60..7850

CDS /*tag= a

FT /product= "signal transduction polypeptide H19G5 splice
FT variant"

XX WO200063381-A1.

FN 26-OCT-2000.

XX 11-APR-2000; 2000WO-US009488.

XX 16-APR-1999; 99US-0129553P.

XX (SCIO-) SCIOS INC.

XX Zeng W, Stanton L, Kong H;

XX WPI; 2001-007013/01.

XX P-PSDB; AAB30569.

XX Novel h19G5 polypeptides capable of regulating signal transduction and
exhibiting kinase activity useful for identifying antibodies to treat
cardiac diseases, and additional mediators of signal transduction.

XX Claim 4; Page 65-68; 81pp; English.

PS The present sequence encodes a splice variant of human in signal
XX transduction polypeptide. The polypeptide is designated H19G5. The
XX protein is capable of regulating signal transduction and exhibits kinase
XX activity. The H19G5 transcript is expressed in the heart. H19G5
XX polypeptides and polynucleotides are useful for preventing or treating a
XX cardiac disease, such as congestive heart failure, dilated congestive
XX cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy,
XX mitral valve disease, aortic valve disease or tricuspid valve disease,
XX angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary,
XX arterial or rheovascular hypertension, arteriosclerosis, atherosclerosis
XX and cardiac tumours in humans. The polypeptide is also useful for
XX detecting the expression of a protein capable of regulating signal
XX transduction or the expression of a protein capable of acting as a donor
XX or acceptor molecule of a phosphate group. The monoclonal antibodies can
XX be used as probes for detecting discrete antigens expressed by tissue or
XX cell samples, and therefore used in humans for localization and
XX monitoring of microbial infection

SQ Sequence 7928 BP; 1535 A; 2644 C; 2488 G; 1261 T; 0 U; 0 Other;

Query Match

99.9%; Score 2999.4; DB 5; Length 7928;

Best Local Similarity 100.0%; Pred. No. 0;			
Matches 3000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
QY	1	CTTACGAGCAGAACTCGGGCCCGCAGGCATACAGGGAGCGAGACATCCCTGGCCCGCCTGAG	60
Db	3443	CTTACGAGCAGAACTCGGGCCCGCAGGCATACAGGGAGCGAGACATCCCTGGCCCGCCTGAG	3502
QY	61	CAACCCGCTGTGTACGGGCTGTGTGACCAAGTTTGTGAGACCCCGCAAGACCCCTCATCTCAT	120
Db	3503	CAACCCGCTGTGTACGGGCTGTGTGACCAAGTTTGTGAGACCCCGCAAGACCCCTCATCTCAT	3562
QY	121	CTGTGAGCTGTCTCATCCGAGAGCTGTGTGACCGCTGTGTACAGGAAAGGGGCTGGTAC	180
Db	3563	CTGTGAGCTGTCTCATCCGAGAGCTGTGTGACCGCTGTGTACAGGAAAGGGGCTGGTAC	3622
QY	181	GGAGGCGGAGGTCAAGGTCTACATCCAGCAGCTGTGTGAGGGGCTGCACCTACCTGCAAG	240
Db	3623	GGAGGCGGAGGTCAAGGTCTACATCCAGCAGCTGTGTGAGGGGCTGCACCTACCTGCAAG	3682
QY	241	CCATGGGCTTTCACCTCGACATAAAGCCCTTAACATCTGTGATGTGCTATCCTGCCCCG	300
Db	3683	CCATGGGCTTTCACCTCGACATAAAGCCCTTAACATCTGTGATGTGCTATCCTGCCCCG	3742
QY	301	GGAGACATTAATAATCTGGACTTTGGCTTTGGCCGAGAAATACACCCGAGCAGAGCTGCA	360
Db	3743	GGAGACATTAATAATCTGGACTTTGGCTTTGGCCGAGAAATACACCCGAGCAGAGCTGCA	3802
QY	361	GTTCAGCAGTACGGCTCCCTGTGATTCGTCTCTCCCGAGATCATCCAGCAGAAACCCCTGT	420
Db	3803	GTTCAGCAGTACGGCTCCCTGTGATTCGTCTCTCCCGAGATCATCCAGCAGAAACCCCTGT	3862
QY	421	GAGGAGACCTCCGACATTTGGGCTATGGGTGTGATCTCTTCACTCAGCCTGACCTGCTC	480
Db	3863	GAGGAGACCTCCGACATTTGGGCTATGGGTGTGATCTCTTCACTCAGCCTGACCTGCTC	3922
QY	481	ATCCCCATTTCCGCGCAGAGTGACCGTGCCACCTCTCTGAACTGTCTGTGAGGGGCGCGT	540
Db	3923	ATCCCCATTTCCGCGCAGAGTGACCGTGCCACCTCTCTGAACTGTCTGTGAGGGGCGCGT	3982
QY	541	GTCATGAGCAGCCCATTTGGTGTGCTGCCACCTCAGCGAAGAGCGCCAAAGACTTTCATCAAGGC	600
Db	3983	GTCATGAGCAGCCCATTTGGTGTGCTGCCACCTCAGCGAAGAGCGCCAAAGACTTTCATCAAGGC	4042
QY	601	TACGCTCGAGAGGCCCTCAGGCCCGGCTAGTGTGGGCCCGAGTGCCTCTCCACCCCTG	660
Db	4043	TACGCTCGAGAGGCCCTCAGGCCCGGCTAGTGTGGGCCCGAGTGCCTCTCCACCCCTG	4102
QY	661	GTTCCTGAAATCCATGCTGTGGAGAGGCGCACTTCATCAACCAAGCAGCAGCTCAAGTT	720
Db	4103	GTTCCTGAAATCCATGCTGTGGAGAGGCGCCACTTCATCAACCAAGCAGCAGCTCAAGTT	4162
QY	721	CCTCCTGGCCGAAAGTCTGCTGGCAGCGTTCCCTGATGAGCTACAAGTCCAATCCTGTGTAT	780
Db	4163	CCTCCTGGCCGAAAGTCTGCTGGCAGCGTTCCCTGATGAGCTACAAGTCCAATCCTGTGTAT	4222
QY	781	GGCTTCATCCCTGAGCTGTGGGGCCCAACCGCAGACCCCTCCCTGGGCTAGCCCG	840
Db	4223	GGCTTCATCCCTGAGCTGTGGGGGGCCCAACCGCAGACCCCTCCCTGGGCTAGCCCG	4282
QY	841	GCACCTTCGAGGAGCACTGGTGGCTCCTCAGTTCTCCTCTCTCTCTCTGACAAACGAGCT	900
Db	4283	GCACCTTCGAGGAGCACTGGTGGCTCCTCAGTTCTCCTCTCTCTCTCTGACAAACGAGCT	4342
QY	901	CGGCCCATTTGCCGGGCTTAAGTCACTGACACCCCTCCCGGCTGACACACTCACCCTGTCT	960
Db	4343	CGGCCCATTTGCCGGGCTTAAGTCACTGACACCCCTCCCGGCTGACACACTCACCCTGTCT	4402
QY	961	GCAACCCCGGGGCTTCTGTGGGCCCTCGGCCAGCCTGTGTGAGGAGCCGAGCCAGTGA	1020
Db	4403	GCAACCCCGGGGCTTCTGTGGGCCCTCGGCCAGCCTGTGTGAGGAGCCGAGCCAGTGA	4462
QY	1021	GGCTCCACCGAGGCCCTCAGCTCCGCTGCACTCCCGAGGGGTGCGGGCCACCCGGCCGC	1080
Db			

Db	4463	GCCTCCACCGAGGCCCCAGCTCCGCTCGCATCTCCGAGGGTCCGGGCCACCGGCCGC	4522
QY	1081	CCAGGGCTGCTGTGCCCGCGGCAACAGGCTCATCCGAGGCTGTCTTCTACCAACAGGCGGGTGA	1140
Db	4523	CCAGGGCTGCTGTGCCCGCGGCAACAGGCTCATCCGAGGCTGTCTTCTACCAACAGGCGGGTGA	4582
QY	1141	GAGCCCTGTAGCAGCGGGCCCTTGGCCCGGGGAGCAGGGGCAACCCGGCCCGGGCGCGCA	1200
Db	4583	GAGCCCTGTAGCAGCGGGCCCTTGGCCCGGGGAGCAGGGGCAACCCGGCCCGGGCGCGCA	4642
QY	1201	CTTCTGAAAGGGGGTAACTTGTGGGGGGCTGTGCCAGGCTGTGCCGAGCCACTGATGGA	1260
Db	4643	CCTGCTGAAAGGGGGTAACTTGTGGGGGGCTGTGCCAGGCTGTGCCGAGCCACTGATGGA	4702
QY	1261	GCACCGGCTGCTGTGAGAGAGAGGCGCGCAGGAGAGGAGCAGGCAACCTCTCTGGCCAAAGC	1320
Db	4703	GCACCGGCTGCTGTGAGAGAGAGGCGCGCAGGAGAGGAGCAGGCAACCTCTCTGGCCAAAGC	4762
QY	1321	CCCTCATTTGAGACTGCGCTCCGGCTGTCTGTGGCAGCCACTTTGGCCCTCTGGCCA	1380
Db	4763	CCCTCATTTGAGACTGCGCTCCGGCTGTCTGTGGCAGCCACTTTGGCCCTCTGGCCA	4822
QY	1381	CAGCACTCCCTGAAACATGACTCTCCGAGCAGCCCGCCCTCTCTCGAGGCTGTGGG	1440
Db	4823	CAGCACTCCCTGAAACATGACTCTCCGAGCAGCCCGCCCTCTCTCGAGGCTGTGGG	4882
QY	1441	TGAGGCACAGCACTGCTTCAAGCCCTCCGGGGGGGGCCCTTATCAGGGACATTTGGGGCA	1500
Db	4883	TGAGGCACAGCACTGCTTCAAGCCCTCCGGGGGGGGCCCTTATCAGGGACATTTGGGGCA	4942
QY	1501	CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGTGTGGCCACCCAGGCACTGTCTCAGGCAGA	1560
Db	4943	CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGTGTGGCCACCCAGGCACTGTCTCAGGCAGA	5002
QY	1561	GAGGCATCTCCCGGACAGCCCTTGGGGGAGCCAGCCCTTTCTGTGCCACCCCAAGCAGGG	1620
Db	5003	GAGGCATCTCCCGGACAGCCCTTGGGGGAGCCAGCCCTTTCTGTGCCACCCCAAGCAGGG	5062
QY	1621	TTCTGCCCCCAGGAGGGCTGCAGCCCCCAGCCAGCAGTTGCCCATGCCCCCTCTGGCTC	1680
Db	5063	TTCTGCCCCCAGGAGGGCTGCAGCCCCCAGCCAGCAGTTGCCCATGCCCCCTCTGGCTC	5122
QY	1681	CTTCTCCTCAGGATCTTGCAAGAGGCCCTTTAGTACCTCAAGCCCTTTCTTGGGACA	1740
Db	5123	CTTCTCCTCAGGATCTTGCAAGAGGCCCTTTAGTACCTCAAGCCCTTTCTTGGGACA	5182
QY	1741	GGCCAGGACACCCCTGCCCCCTGCCAAGCAAGCCCCCATTTGGACTCTTAAGATGGGGCC	1800
Db	5183	GGCCAGGACACCCCTGCCCCCTGCCAAGCAAGCCCCCATTTGGACTCTTAAGATGGGGCC	5242
QY	1801	TGGAGACATCTCTTCTTGGGAGGCCAAACCCCGGCCCTGCAGTTCCCAAGGCTCAGC	1860
Db	5243	TGGAGACATCTCTTCTTGGGAGGCCAAACCCCGGCCCTGCAGTTCCCAAGGCTCAGC	5302
QY	1861	CTTCCAGGCGAGCTTTCCCAAGTGAAGTCTCTCAAGGTTGGGCTCTCTCCAGGTTGGGCAC	1920
Db	5303	CTTCCAGGCGAGCTTTCCCAAGTGAAGTCTCTCAAGGTTGGGCTCTCTCCAGGTTGGGCAC	5362
QY	1921	AGAGCTTGGCCCTCTCTGGATTCGGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA	1980
Db	5363	AGAGCTTGGCCCTCTCTGGATTCGGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA	5422
QY	1981	CTTCCACACCCCTTTCAGCGGCTCAGGAAACAGGCGACCATGCGCAAGTTCTCTCCCTGG	2040
Db	5423	CTTCCACACCCCTTTCAGCGGCTCAGGAAACAGGCGACCATGCGCAAGTTCTCTCCCTGG	5482
QY	2041	TGCTCGCGGGGCTACGACGCGGTGGCTGTATGGCACTTTGCTTTTGTGGAGATGTC	2100
Db	5483	TGCTCGCGGGGCTACGACGCGGTGGCTGTATGGCACTTTGCTTTTGTGGAGATGTC	5542
QY	2101	AGGGGGCATGTCTGGGGCAGGGGCCCATGTGTGGCCAGGATAGCTGGGGCTGTCCCATC	2160
Db	5543	AGGGGGCATGTCTGGGGCAGGGGCCCATGTGTGGCCAGGATAGCTGGGGCTGTGTCCCATC	5602

QY 61 CCACCCGCTGGTCA CCGGGCTCTCTGGACCAAGTTTGAGACCCCGAAGACCCCTCATCTCAT 120
DB |||||
759 CCACCCGCTGGTCA CCGGGCTCTCTGGACCAAGTTTGAGACCCCGAAGACCCCTCATCTCAT 818
QY 121 CCTGGAGCTGTCTCAT CCGAGGAGCTGTGACCCCTCTGTATACAGAAAGGGGCTGGTGA 180
DB |||||
819 CTTGGAGCTGTCTCAT CCGAGGAGCTGTGACCCCTGTACAGAAAGGGGCTGGTGA 878
QY 181 GGAGGCCGAGGTCAAGGCTTACATCCAGCAGCTGGTGGAGGGGCTGCACTACCTGSCACAG 240
DB |||||
879 GGAGGCCGAGGTCAAGGCTTACATCCAGCAGCTGGTGGAGGGGCTGCACTACCTGSCACAG 938
QY 241 CCATGGGCTTCTGCA CCGTGGAATAAAGCCTCTTAACATCTGTATGATGCTGCTGCCG 300
DB |||||
939 CCATGGGCTTCTGCA CCGTGGAATAAAGCCTCTTAACATCTGTATGATGCTGCTGCCG 998
QY 301 GGAGACATTAATATCTCGACTTTGGCTTTGGCCAGAAATCAACCCAGCAGAGCTGCA 360
DB |||||
999 GGAGACATTAATATCTCGACTTTGGCTTTGGCCAGAAATCAACCCAGCAGAGCTGCA 1058
QY 361 GTTCAGCCAGTACCGCTCCCTCGAGTTGCTCTCTCCCGAGATCATCCAGCAGAAACCTGT 420
DB |||||
1059 GTTCAGCCAGTACCGCTCCCTCGAGTTGCTCTCTCCCGAGATCATCCAGCAGAAACCTGT 1118
QY 421 GAGCGAGCCTCCGACATTTGGGCCATGGGTGTCACTCTCTACCTCAGCCTGACCTGCTC 480
DB |||||
1119 GAGCGAGCCTCCGACATTTGGGCCATGGGTGTCACTCTCTCTACCTCAGCCTGACCTGCTC 1178
QY 481 ATCCCCATTTGCCCGCAGAGTGACCGGTGCCACCTCTCTGAACGTCTCTGGAGGGGCGCGT 540
DB |||||
1179 ATCCCCATTTGCCCGCAGAGTGACCGGTGCCACCTCTCTGAACGTCTCTGGAGGGGCGCGT 1238
QY 541 GTCATGAGCAGCCCCATTTGGTGCCCACTCAGCGAAGACGCCAAAGACTTTCATCAAGGC 600
DB |||||
1239 GTCATGAGCAGCCCCATTTGGTGCCCACTCAGCGAAGACGCCAAAGACTTTCATCAAGGC 1298
QY 601 TACCTCGCAGAGCCCTCAGGCCCGGCTAGTGGCGGCCAGTGCCTCTCCACCCCTG 660
DB |||||
1299 TACCTCGCAGAGCCCTCAGGCCCGGCTAGTGGCGGCCAGTGCCTCTCCACCCCTG 1358
QY 661 GTTCCTGAATTCATGCTCGGAGGAGGCGCCACTTCATCAACACCAAGCAGCTCAAGTT 720
DB |||||
1359 GTTCCTGAATTCATGCTCGGAGGAGGCGCCACTTCATCAACACCAAGCAGCTCAAGTT 1418
QY 721 CTTCTGGCCGGAAGTCTGGCAGCGTTCCCTGTATGAGCTACAAATTCATCTGGTGAT 780
DB |||||
1419 CTTCTGGCCGGAAGTCTGGCAGCGTTCCCTGTATGAGCTACAAATTCATCTGGTGAT 1478
QY 781 GCGCTCCATCCCTGAGCTGTGGGGGCCACCGCAGACCCCTCTCGGCGGTAGCCG 840
DB |||||
1479 GCGCTCCATCCCTGAGCTGTGGGGGCCACCGCAGACCCCTCTCGGCGGTAGCCG 1538
QY 841 GCACCTTGAGGACACTGGTGCTCTCTCAGTTCTCTCTCTCTCTCTCTCTGACAAACGAGCT 900
DB |||||
1539 GCACCTTGAGGACACTGGTGCTCTCTCAGTTCTCTCTCTCTCTCTCTCTGACAAACGAGCT 1598
QY 901 CGCCCAATTTGCCGGGCTAAGTCACTGACCCCTCCCGGTGACACACTCAGCACTGCT 960
DB |||||
1599 CGCCCAATTTGCCGGGCTAAGTCACTGACCCCTCCCGGTGACACACTCAGCACTGCT 1658
QY 961 GCACCCCGGGGCTTCTCGGGGCTCTCGGCGCCCTGAGGAGGCTGCTGAGGAGGCGAGGCGAGTGA 1020
DB |||||
1659 GCACCCCGGGGCTTCTCGGGGCTCTCGGCGCCCTGAGGAGGCTGCTGAGGAGGCGAGGCGAGTGA 1718
QY 1021 GCGCTTCAACGAGGCCAGCTCCGCTGTATCTCCGAGGCTGCGGGGCCACCGGCGGC 1080
DB |||||
1719 GCGCTTCAACGAGGCCAGCTCCGCTGTATCTCCCGAGGCTGCGGGGCCACCGGCGGC 1778
QY 1081 CCAGGCTGTGCTCCCGGCACAGCTCATCCGACGCTGTCTACCAACGAGCGGCTGA 1140
DB |||||
1779 CCAGGCTGTGCTCCCGGCACAGCTCATCCGACGCTGTCTACCAACGAGCGGCTGA 1838

QY 1141 GAGCCTTGAGCACGGGCGCTTGCGCCCGGGAGCAGCGCGCACCCGCGCGCGCGGCA 1200
DB |||||
1839 GAGCCTTGAGCACGGGCGCTTGCGCCCGGGAGCAGCGCGCACCCGCGCGCGCGGCA 1898
QY 1201 CTTGCTGAAGGGGCGGTACATTTGCGGGGCGCTGCCAGGCTTGCAGGACCTGATGGA 1260
DB |||||
1899 CTTGCTGAAGGGGCGGTACATTTGCGGGGCGCTGCCAGGCTTGCAGGACCTGATGGA 1958
QY 1261 GCACCGGCTGCTGGAGGAGGCGCGCCAGGGAGGAGCAGGCGCACCTCTGCGCAAGC 1320
DB |||||
1959 GCACCGGCTGCTGGAGGAGGCGCGCCAGGGAGGAGCAGGCGCACCTCTGCGCAAGC 2018
QY 1321 CCCCCTCATTTGAGACTGCTCCGCTTGCCTCTCTGGCACCCACTTTGGCCCCCTGCGCA 1380
DB |||||
2019 CCCCCTCATTTGAGACTGCTCCGCTTGCCTCTCTGGCACCCACTTTGGCCCCCTGCGCA 2078
QY 1381 CAGGCACTCCCTGGAACTGCTCTCGAGCA CCCCCCGCCCTCTCTCGAGGCTGCGG 1440
DB |||||
2079 CAGGCACTCCCTGGAACTGCTCTCGAGCA CCCCCCGCCCTCTCTCGAGGCTGCGG 2138
QY 1441 TGAGGCAACAGCACTGCTTTCAGCCCCCTCCGGGGGGGCCCCCTATCAGGACATGGGGCA 1500
DB |||||
2139 TGAGGCAACAGCACTGCTTTCAGCCCCCTCCGGGGGGGCCCCCTATCAGGACATGGGGCA 2198
QY 1501 CCTCAGGGCTCCAAGCAGCTTCCATCCATCTGTGGGCCACCCAGGCACTGCTCAGCGAGA 1560
DB |||||
2199 CCTCAGGGCTCCAAGCAGCTTCCATCCATCTGTGGGCCACCCAGGCACTGCTCAGCGAGA 2258
QY 1561 GAGGCCATCCCGGACAGCCTTTGGGGGAGCAGCAGCCCTTTCTGCAACCCCAAGCAGGG 1620
DB |||||
2259 GAGGCCATCCCGGACAGCCTTTGGGGGAGCAGCAGCCCTTTCTGCAACCCCAAGCAGGG 2318
QY 1621 TTCTGCCCCCAGAGGGGTGACGCCCCCAGCAGAGTTGCCCATGCCCCCTCTCTCGGCTC 1680
DB |||||
2319 TTCTGCCCCCAGAGGGGTGACGCCCCCAGCAGAGTTGCCCATGCCCCCTCTCTCGGCTC 2378
QY 1681 CTTTCCCTCAGAGATCTTGCAAGAGGCCCCCTTAGTACCCTCAAGCCCCCTTTTGGGACA 1740
DB |||||
2379 CTTTCCCTCAGAGATCTTGCAAGAGGCCCCCTTAGTACCCTCAAGCCCCCTTTTGGGACA 2438
QY 1741 GCCCCAGGACCCCCCTGCCCCCTGCCAAAGCAAGCCCCCATTTGGACTCTAAGATGGGGCC 1800
DB |||||
2439 GCCCCAGGACCCCCCTGCCCCCTGCCAAAGCAAGCCCCCATTTGGACTCTAAGATGGGGCC 2498
QY 1801 TGGAGACATCTCTCTTCTTCTGGAGGCGCAAAACCCCGGCCCTTCAGTTCCCCAGGGTCAGC 1860
DB |||||
2499 TGGAGACATCTCTCTTCTTCTGGAGGCGCAAAACCCCGGCCCTTCAGTTCCCCAGGGTCAGC 2558
QY 1861 CTTCCAGGCGAGCTCTTCCCAAGTGAAGTCCCTCAGGGTGGGCTCTCTCCAGGTGGGCAC 1920
DB |||||
2559 CTTCCAGGCGAGCTCTTCCCAAGTGAAGTCCCTCAGGGTGGGCTCTCTCCAGGTGGGCAC 2618
QY 1921 AGAGCTGCGGCCCTCCCTGGATGCGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 1980
DB |||||
2619 AGAGCTGCGGCCCTCCCTGGATGCGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 2678
QY 1981 CTTCCACACCCACTTTGCAAGCGGCTCAGGAAACAGCGACCATGCGCAAGTTCTCCCTGGG 2040
DB |||||
2679 CTTCCACACCCACTTTGCAAGCGGCTCAGGAAACAGCGACCATGCGCAAGTTCTCCCTGGG 2738
QY 2041 TGGTCGCGGGGGCTACGAGGCGGTGGCTGCTATGGCA CTTTGGCTTTGGTGGAGATGC 2100
DB |||||
2739 TGGTCGCGGGGGCTACGAGGCGGTGGCTGCTATGGCA CTTTGGCTTTGGTGGAGATGC 2798
QY 2101 AGGGGCACTGCTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGTCCAGTC 2160
DB |||||
2799 AGGGGCACTGCTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGTCCAGTC 2858
QY 2161 GGAAGAGAGGAGCAAGAGGAGGCGCGGCTGAGTTCCAGTCTGAGAGGAGCAGCAGAGGC 2220
DB |||||
2859 GGAAGAGAGGAGCAAGAGGAGGCGCGGCTGAGTTCCAGTCTGAGAGGAGCAGCAGAGGC 2918
QY 2221 CAGGGCTGAGAGCCACTGTCGCCAGGTCAGTGCAAGGCCCTGTGCTGAGGTCGGCAGGGC 2280


```
Db 2919 CAGGGCTGAGAGCCCACTGCCCGAGGTGAGTGAAGGCTGTGCTGAGGTGCGGAGGGC 2978
Qy 2281 TCCACACGAGAGCTCTCCAGAGCCACCCCATGCGGAGGACATCGGGCAGGTCTCCCTGGT 2340
Db 2979 TCCACACGAGAGCTCTCCAGAGCCACCCCATGCGGAGGACATCGGGCAGGTCTCCCTGGT 3038
Qy 2341 GCAGATCCGGAGCCTGTCAAGTGATGCGGAGGCGCGGACACAATATCCCTGGACATTC 2400
Db 3039 GCAGATCCGGAGCCTGTCAAGTGATGCGGAGGCGCGGACACAATATCCCTGGACATTC 3098
Qy 2401 CGAGGTGAGACCCCGCTACCTCAACCTCTCAGAGCTGTACGATATCAAGTACCTCCCAT 2460
Db 3099 CGAGGTGAGACCCCGCTACCTCAACCTCTCAGAGCTGTACGATATCAAGTACCTCCCAT 3158
Qy 2461 CGAGTTTATGATCTTCAGGAAGTTCCTGAGAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCCAT 2520
Db 3159 CGAGTTTATGATCTTCAGGAAGTTCCTGAGAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCCAT 3218
Qy 2521 GGCTGAGGAGAGCTGGCCGAGTTCCCGGAGCCCACTGCGCCCTGGCCAGTGAAGTGG 2580
Db 3219 GGCTGAGGAGAGCTGGCCGAGTTCCCGGAGCCCACTGCGCCCTGGCCAGTGAAGTGG 3278
Qy 2581 CCCCACGCGAGCCTGAGAGTACAGAGAGTCAAGAGTGTGACGCGCTGCTGGGAGA 2640
Db 3279 CCCCACGCGAGCCTGAGAGTACAGAGAGTCAAGAGTGTGACGCGCTGCTGGGAGA 3338
Qy 2641 GGCTGCGTGGGAGGAGCGCAAGTGTCTTCGCCCTCACGAGCCTCTTCCACTTCCC 2700
Db 3339 GGCTGCGTGGGAGGAGCGCAAGTGTCTTCGCCCTCACGAGCCTCTTCCACTTCCC 3398
Qy 2701 TGGAGGACCTGCGCTGATGAGCTGACAGAGTGGGGTGGCTGAGAGAGTGAAGGC 2760
Db 3399 TGGAGGACCTGCGCTGATGAGCTGACAGAGTGGGGTGGCTGAGAGAGTGAAGGC 3458
Qy 2761 CTGCGTGAGACATCTCCCGATCTCAGAGGAGCGGAGGCTCTGGAAGAGGAGG 2820
Db 3459 CTGCGTGAGACATCTCCCGATCTCAGAGGAGCGGAGGCTCTGGAAGAGGAGG 3518
Qy 2821 GCGCCCGAGGAGGAGCGGCTTGTCTTCTCCGCTCTCAGGTCTGAAGAGCTGGGA 2880
Db 3519 GCGCCCGAGGAGGAGCGGCTTGTCTTCTCCGCTCTCAGGTCTGAAGAGCTGGGA 3578
Qy 2881 CCGAGCGCGGACATCTCAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCAGT 2940
Db 3579 CCGAGCGCGGACATCTCAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCAGT 3638
Qy 2941 GACACTGCGCTGCGAGGTGTGAGCCAGCAGCTGCGGAGCCACCTGGAGCAAGAGCG 3000
Db 3639 GACACTGCGCTGCGAGGTGTGAGCCAGCAGCTGCGGAGCCACCTGGAGCAAGAGCG 3698
```

RESULT 7

AAL43908
ID AAL43908 standard; DNA; 5207 BP.

XX AAL43908;

AC AAL43908;

DT 19-SEP-2002 (first entry)

XX Human kinase protein coding sequence 1.

XX Human; gene; ds; gene therapy; chromosome 1; kinase protein;

KW myosin light chain kinase subfamily; kinase protein-mediated disease;

XX transgenic animal.

XX Homo sapiens.

XX Key Location/Qualifiers

PH

FT CDS

109..5106

/**tag=a
/product= "Human kinase protein"

WO200240683-A2.

23-MAY-2002.

22-OCT-2001; 2001WO-US032616.

14-NOV-2000; 2000US-00711134.

17-MAY-2001; 2001US-00858664.

(PEKE) PE CORP NY.

Wei M, Ketchum K, Di Francesco V, Beasley EM;
P-PSDB; AAO15372.

WPI; 2002-500223/53.

New kinase proteins related to myosin light chain kinase subfamily and
encoding polynucleotide, useful for diagnosing, treating disease or
condition mediated by the kinase protein and for identifying modulators.

Claim 23; Fig 1; 96pp; English.

The invention comprises the amino acid and coding sequences (located on
chromosome 1) of a human kinase protein that is related to the myosin
light chain kinase subfamily. The human kinase DNA and protein sequences
of the invention are useful for identifying agents that modulate the
activity of the human kinase protein. Kinase-modulating agents are useful
for treating a disease or condition mediated by a human kinase protein.
The human kinase DNA sequences can be used to produce transgenic animals
which are useful for studying the function of kinase proteins and
identifying/evaluating modulators of kinase protein activity. The present
DNA sequence encodes the human kinase protein of the invention

Sequence 5207 BP; 988 A; 1758 C; 1612 G; 849 T; 0 U; 0 Other;

Query Match 99.9%; Score 2997.8; DB 6; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTACGAGCAGAACTCGGGCCAGGCATACAGAGGAGGAGAGACATCTGGCGCGGCTGAG 60
Db 699 CCTACGAGCAGAACTCGGGCCAGGCATACAGAGGAGGAGAGACATCTGGCGCGGCTGAG 758
Qy 61 CCACCCGCTGTCTACGCGGGCTGTGGACCAAGTTTGTAGACCGCAAGACCCCTCATCCTCAT 120
Db 759 CCACCCGCTGTCTACGCGGGCTGTGGACCAAGTTTGTAGACCGCAAGACCCCTCATCCTCAT 818
Qy 121 CCTGGAGCTGTGCTCATCCGAGGAGCTGTGGACCGCTGTACAGGAGGCGGTGGTGAC 180
Db 819 CCTGGAGCTGTGCTCATCCGAGGAGCTGTGGACCGCTGTACAGGAGGCGGTGGTGAC 878
Qy 181 GGAGCGCGAGGTCAAGGTCTACATCCAGAGCTGTGGTGGAGGGGCTGCATCTGACACAG 240
Db 879 GGAGCGCGAGGTCAAGGTCTACATCCAGAGCTGTGGTGGAGGGGCTGCATCTGACACAG 938
Qy 241 CCATGGGCTTCTCCACCTGGACATAAAGCCCTCTAATCATCTGTATGGTGCATCTGCGCG 300
Db 939 CCATGGGCTTCTCCACCTGGACATAAAGCCCTCTAATCATCTGTATGGTGCATCTGCGCG 998
Qy 301 GGAAGACATTAATAATCTGCGACTTTGGCTTTGGCCAGAAACATCAACCCAGCAGAGCTGCA 360
Db 999 GGAAGACATTAATAATCTGCGACTTTGGCTTTGGCCAGAAACATCAACCCAGCAGAGCTGCA 1058
Qy 361 GTTCAGCAGTACCGCTCCCTGTGATTCGTCTCCCGAGATCATCCAGCAGAAACCTGT 420
Db 1059 GTTCAGCAGTACCGCTCCCTGTGATTCGTCTCCCGAGATCATCCAGCAGAAACCTGT 1118
Qy 421 GAGCGAGCCTCCGACATTTGGGCGCATGGGTGTCTCTCTACCTCAGCCTGACCTGCTC 480

Db 1119 GAGCGAAGCCTCCGACATTTGGGCCATGGGTGTATCTCTACCTCAGCTGACCTGTCTC 1178
QY 481 ATCCCAATTGCGGCGAGGTGAACCGTGCACACCTCTCTGAACGTCCTGGAGGGCGCGT 540
Db 1179 ATCCCAATTGCGGCGAGGTGACCGTGCACACCTCTCTGAACGTCCTGGAGGGCGCGT 1238
QY 541 GTCATGGAGCAGCCCATGGCTGCCACCTCAGCGAAGACGCGCAAGACTTCATCAAGGC 600
Db 1239 GTCATGGAGCAGCCCATGGCTGCCACCTCAGCGAAGACGCGCAAGACTTCATCAAGGC 1298
QY 601 TACGCTGCAGAGAGCCCTCAGGCCGCGCTAGTGGGCGCCAGTGCCCTCCACCCCTG 660
Db 1299 TACGCTGCAGAGAGCCCTCAGGCCGCGCTAGTGGGCGCCAGTGCCCTCCACCCCTG 1358
QY 661 GTTCTGAAATCCATCCCTGGGAGAGGCCACATTCATCAACAAAGCAGCTCAAGTT 720
Db 1359 GTTCTGAAATCCATCCCTGGGAGAGGCCACATTCATCAACAAAGCAGCTCAAGTT 1418
QY 721 CCTCTGGCCCGAAGTCCGCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTGGTGAT 780
Db 1419 CCTCTGGCCCGAAGTCCGCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTGGTGAT 1478
QY 781 GCGCTCCATCCCTGAGCTGCTGCGGGGCCACCCGACAGCCCTCCCTCGGCGTAGCCG 840
Db 1479 GCGCTCCATCCCTGAGCTGCTGCGGGGCCACCCGACAGCCCTCCCTCGGCGTAGCCG 1538
QY 841 GCACTCTGCGAGGCACTGCTGGCTCCTCCAGTTCTCTCTCTCTCTCTGACAAAGAGCT 900
Db 1539 GCACTCTGCGAGGCACTGCTGGCTCCTCCAGTTCTCTCTCTCTCTCTGACAAAGAGCT 1598
QY 901 CGCCCCATTTGCGGGCTAGTCACTGCGCACCTTCCCGGTGACACACTCACCACTGCT 960
Db 1599 CGCCCCATTTGCGGGCTAGTCACTGCGCACCTTCCCGGTGACACACTCACCACTGCT 1658
QY 961 GCACCCCGGGCTTCTGCGGCCCTCGGCCAGCTGCTGAGGAAGCCGAGGCCAGTGA 1020
Db 1659 GCACCCCGGGCTTCTGCGGCCCTCGGCCAGCTGCTGAGGAAGCCGAGGCCAGTGA 1718
QY 1021 GCGTCTCACCGAGGCCAGCTCGCTGCAATCTCCGAGGTCGCGGCCACCGGCCG 1080
Db 1719 GCGTCTCACCGAGGCCAGCTCGGCCCTGCAATCTCCGAGGTCGCGGCCACCGGCCG 1778
QY 1081 CCAGGCTGCGTGCCTCGGCCAGCTGCTGCAATCTCCGAGGTCGCGGCCACCGGCCG 1140
Db 1779 CCAGGCTGCGTGCCTCGGCCAGCTGCTGCAATCTCCGAGGTCGCGGCCACCGGCCG 1838
QY 1141 GAGCCTTGAGCAGCGGGCCCTTGGCCCGGGGAGCAGCGGCACCCGCGCCGCGCGGCA 1200
Db 1839 GAGCCTTGAGCAGCGGGCCCTTGGCCCGGGGAGCAGCGGCACCCGCGCCGCGCGGCA 1898
QY 1201 CTTGCTGAAGGGCGGTACATTTGCGGGGCGCTGCCAGGCTGCGGAGCCACTGATGGA 1260
Db 1899 CTTGCTGAAGGGCGGTACATTTGCGGGGCGCTGCCAGGCTGCGGAGCCACTGATGGA 1958
QY 1261 GCACCGGTGCTGGAGAGAGGCGCGCCAGGAGGAGCAGGCCACCTCTCTGCGCAAGC 1320
Db 1959 GCACCGGTGCTGGAGAGAGGCGCGCCAGGAGGAGCAGGCCACCTCTCTGCGCAAGC 2018
QY 1321 CCCTCTATTGAGACTGSCCTCGGCTGCTGCTGCGCACCACTTGGCCCTCGGCCA 1380
Db 2019 CCCTCTATTGAGACTGSCCTCGGCTGCTGCTGCGCACCACTTGGCCCTCGGCCA 2078
QY 1381 CAGCCACTCCCTGGAACATGACTCTCGAGCAACCCCGGCCCTCTCTGAGGAGCTGCGG 1440
Db 2079 CAGCCACTCCCTGGAACATGACTCTCGAGCAACCCCGGCCCTCTCTGAGGAGCTGCGG 2138
QY 1441 TGAGGACACAGCACTGCTTACGCCCTCCGCGGGGGCCCTTATCAGGGAATGAGGCA 1500
Db 2139 TGAGGACACAGCACTGCTTACGCCCTCCGCGGGGGGGCCCTTATCAGGGAATGAGGCA 2198
QY 1501 CCTCAGGGCTCCAGGAGCTTCCATCTGCTGAGCCACCCAGGCACTGCTCAGCCAGA 1560
Db 2199 CCTCAGGGCTCCAGGAGCTTCCATCTGCTGAGCCACCCAGGCACTGCTCAGCCAGA 2258

QY 1561 GAGGCCATCCCGGACAGCCCTTGGGGGCGAGCCAGCCCTTTCTTGCCACCCCAAGCAGGG 1620
Db 2259 GAGGCCATCCCGGACAGCCCTTGGGGGCGAGCCAGCCCTTTCTTGCCACCCCAAGCAGGG 2318
QY 1621 TTCTGCCCCCCCAGAGAGGCTGAGCCGCCCAAGCAGTGTGCCCATGCCCTCTCTGGCTC 1680
Db 2319 TTCTGCCCCCCCAGAGAGGCTGAGCCGCCCAAGCAGTGTGCCCATGCCCTCTCTGGCTC 2378
QY 1681 CTTCCCTCCAGGATTTGCAAGAGGCCCCCTTAGTACCCTCAAGCCCCCTTTCTTGGGACA 1740
Db 2379 CTTCCCTCCAGGATTTGCAAGAGGCCCCCTTAGTACCCTCAAGCCCCCTTTCTTGGGACA 2438
QY 1741 GCCCCAGGCAACCCCTGCGCCCTGCAAGCAAGCCCCCAAGTGTGAGTCTAAGATGGGGCC 1800
Db 2439 GCCCCAGGCAACCCCTGCGCCCTGCAAGCAAGCCCCCAAGTGTGAGTCTAAGATGGGGCC 2498
QY 1801 TGAGAGCATCTCTCTTCTGCGAGGCCAAACCCGCCCCCTGAGATTTCCCAGGGTCAGC 1860
Db 2499 TGAGAGCATCTCTCTTCTGCGAGGCCAAACCCGCCCCCTGAGATTTCCCAGGGTCAGC 2558
QY 1861 CTCCAGGCGAGCTCTTCCCAAGTGAAGTCCCTCAGGCTGGGCTCTCTCCAGGTGGGAC 1920
Db 2559 CTCCAGGCGAGCTCTTCCCAAGTGAAGTCCCTCAGGCTGGGCTCTCTCCAGGTGGGAC 2618
QY 1921 AGAGCCTTGGGCCCTCCTGATGCGAGGCTGAGCCAGGAGGCTGAGGATCTGTCCGA 1980
Db 2619 AGAGCCTTGGGCCCTCCTGATGCGAGGCTGAGCCAGGAGGCTGAGGATCTGTCCGA 2678
QY 1981 CTCACACCCACCTTGGAAGCGGCTCAGGAACAGGCGCACATGCGCAAGTTCTCCCTGGG 2040
Db 2679 CTCACACCCACCTTGGAAGCGGCTCAGGAACAGGCGCACATGCGCAAGTTCTCCCTGGG 2738
QY 2041 TGCTCGCGGGGCTACGAGGCGTGGCTGGCTATGCGACCTTTGCTTGGTGGAGATGC 2100
Db 2739 TGCTCGCGGGGCTACGAGGCGTGGCTGGCTATGCGACCTTTGCTTGGTGGAGATGC 2798
QY 2101 AGGGGCGATGCTGGGGCAGGGGCCATGTGGGCGAGGATAGCTGGGCTGTGTCCAGTC 2160
Db 2799 AGGGGCGATGCTGGGGCAGGGGCCATGTGGGCGAGGATAGCTGGGCTGTGTCCAGTC 2858
QY 2161 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Db 2859 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2918
QY 2221 CAGGGCTGAGAGCCACCTGCCCCAGTCAAGGCTGTGCTGAGGCTGCGGAGGAGG 2280
Db 2919 CAGGGCTGAGAGCCACCTGCCCCAGTCAAGGCTGTGCTGAGGCTGCGGAGGAGG 2978
QY 2281 TCCACAGGAGCTCTCCAGAGCCACCCCATGGGAGGACATCGGGCAGGTCCTCCCTGCT 2340
Db 2979 TCCACAGGAGCTCTCCAGAGCCACCCCATGGGAGGACATCGGGCAGGTCCTCCCTGCT 3038
QY 2341 GCAGATCCGGGAGCTGTGAGTGATGCGAGGGCGCGACACAATATCCCTGGACATTTTC 2400
Db 3039 GCAGATCCGGGAGCTGTGAGTGATGCGAGGGCGCGCGACACAATATCCCTGGACATTTTC 3098
QY 2401 CGAGGTGAGACCCCGCTACTCAACTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
Db 3099 CGAGGTGAGACCCCGCTACTCAACTCTCAGACCTGTACGATATCAAGTACCTCCCAT 3158
QY 2461 CGAGTTTATGATTTTCAGGAAAGTCCCAAGTCCCTCAGCCAGAGCCGCCCTCCCCCAT 2520
Db 3159 CGAGTTTATGATTTTCAGGAAAGTCCCAAGTCCCTCAGCCAGAGCCGCCCTCCCCCAT 3218
QY 2521 GCTCAGGAGGAGCTGCGCGAGTTCCCGAGGCCACGCTGGGCCCTGCGCAGGTGAATGGG 2580
Db 3219 GCTCAGGAGGAGCTGCGCGAGTTCCCGAGGCCACGCTGGGCCCTGCGCAGGTGAATGGG 3278
QY 2581 CCCCCACGAGGCTGAGAGTACACAGAGGAGTCAAGAGATGAGGAGTGTGGACGCTGTGGCAGA 2640
Db 3279 CCCCCACGAGGCTGAGAGTACACAGAGGAGTCAAGAGGATGAGGAGTGTGGACGCTGTGGCAGA 3338

Qy 2641 GGCTGCGTGGGAGGAGCGCAAGTGGTCTCGCGCTCAGCGAGCCTCTTCCACTTCCC 2700
Dy 3339 GGCTGCGTGGGAGGAGCGCAAGTGGTCTCGCGCTCAGCGAGCCTCTTCCACTTCCC 3398
Qy 2701 TGGGAGGACACCTGCGCGTGGATGAGCCTGACAGAGCTGGGGCTGGCGTGAGAGAGTGAAGGC 2760
Dy 3399 TGGGAGGACACCTGCGCGTGGATGAGCCTGACAGAGCTGGGGCTGGCGTGAGAGAGTGAAGGC 3458
Qy 2761 CTCGCTGGAGCACATCTCCCGGATCTGAAGGCGAGCGCGGAAGTCTGGAGAGAGGAGG 2820
Dy 3459 CTCGCTGGAGCACATCTCCCGGATCTGAAGGCGAGCGCGGAAGTCTGGAGAGAGGAGG 3518
Qy 2821 GCGCCCGAGGAGAGCGAGCGCTTGGCTTCTCGCGCTCTCAGCTCTGAAGAGCTGCGA 2880
Dy 3519 GCGCCCGAGGAGAGCGAGCGCTTGGCTTCTCGCGCTCTCAGCTCTGAAGAGCTGCGA 3578
Qy 2881 CCGAGCGCGGACATCTTCCAGGAGCTCTCAGATGAGACTGTGGTCTGGGCGGACGTCAGT 2940
Dy 3579 CCGAGCGCGGACATCTTCCAGGAGCTCTCAGATGAGACTGTGGTCTGGGCGGACGTCAGT 3638
Qy 2941 GACACTGGCTGCGAGGTGTGAGCGCCAGCGAGCTGCCAGGCCACCTGGAGCAAGAGCG 3000
Dy 3639 GACACTGGCTGCGAGGTGTGAGCGCCAGCGAGCTGCCAGGCCACCTGGAGCAAGAGCG 3698
Qy 3001 A 3001
Dy 3699 A 3699
RESULT 8
AAD38865
ID AAD38865 standard; cdna; 5454 BP.
AC AAD38865;
XX
DT 23-SEP-2002 (first entry)
XX
DE Human kinase (PKIN)-22; cdna.
XX
KW Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
KW development; hepatitis; cardiovascular; hypertension; drug screening;
KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
KW hypercholesterolaemia; obesity; gene therapy; cytostatic; anti-hiv;
KW neuroprotective; hepatotropic; hypotensive; cardiac; nephrotropic;
KW hyperlipidaemia; enzyme; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 373..5370
FT /tag= a
FT /product= "Human kinase (PKIN)-22"
XX
PN WO200233099-A2.
XX
PD 25-APR-2002.
XX
PF 20-OCT-2001; 2001WO-US047728.
XX
PR 20-OCT-2000; 2000US-0242410P.
PR 27-OCT-2000; 2000US-0244068P.
PR 03-NOV-2000; 2000US-0245708P.
PR 09-NOV-2000; 2000US-0247672P.
PR 16-NOV-2000; 2000US-0249565P.
PR 22-NOV-2000; 2000US-0252730P.
PR 01-DEC-2000; 2000US-0250807P.
XX
XX (INCY-) INCYTE GENOMICS INC.
XX
PI Gururajan R, Baughn MR, Wallia NK, Elliott VS, Xu Y, Arvizu C;

PI Yao MG, Ramkumar J, Ding L, Tang YT, Hafalia AJA, Nguyen DB;
PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG;
PI Thangavelu K, Lu DAM, Borowsky ML, Thornton M, Swannaker A;
XX WPI; 2002-454603/48.
DR P-PSDB; AAE24151.
XX
PS New human kinase polypeptide, for diagnosing, preventing and treating
PT cancer, immune system disorders, growth and development disorders,
PT cardiovascular disorders and lipid disorders.
XX
PS Claim 5; Page 209-210; 210pp; English.
XX
CC The invention relates human kinases (PKIN) and their corresponding
CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
CC treating and preventing cancer, an immune system disorder (e.g., acquired
CC immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,
CC atherosclerosis, multiple sclerosis, psoriasis), disorders affecting
CC growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),
CC cardiovascular disorder (e.g., hypertension, myocardial infarction,
CC Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,
CC Gaucher's disease, Niemann-Pick's disease, hypercholesterolaemia,
CC hyperlipidaemia, obesity), and for assessing the effects of exogenous
CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a
CC condition or a disease associated with the expression of PKIN in a
CC biological sample. A composition comprising PKIN or an agonist or
CC antagonist of PKIN is useful for treating a disease or condition
CC associated with decreased or increased expression of functional PKIN.
CC PKIN is useful in a number of drug screening techniques and to analyse
CC the proteome of a tissue or cell type. PKIN DNA is useful for creating
CC knockin humanised animals or transgenic animals to model human disease,
CC and in somatic or germline gene therapy. The present sequence is human
CC PKIN cdna
XX
SQ Sequence 5454 BP; 1019 A; 1853 C; 1686 G; 896 T; 0 U; 0 Other;
Query Match 99.8%; Score 2996.2; DB 6; Length 5454;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2998; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CCTACGAGCAGAACTCGGGCCAGGATACAGGGAGGAGAGATCTTGGCGGCTCAG 60
Dy 963 CCTACGAGCAGAACTCGGGCCAGGATACAGGGAGGAGAGATCTTGGCGGCTCAG 1022
Qy 61 CCACCCCGTGTCAACGGGGCTGTGGACCAAGTTTGGAGACCCGCAAGACCTCATCTCAT 120
Dy 1023 CCACCCCGTGTCAACGGGGCTGTGGACCAAGTTTGGAGACCCGCAAGACCTCATCTCAT 1082
Qy 121 CTGTGAGCTGTCTCATCCGAGGAGCTGTGGACCGCTGTACAGGAAGGCGTGGTGCAC 180
Dy 1083 CTGTGAGCTGTCTCATCCGAGGAGCTGTGGACCGCTGTACAGGAAGGCGTGGTGCAC 1142
Qy 181 GGAGCGGAGGTCAAGGTCTACATCCAGAGCTGTGTGGAGGGCTGTACTACCTGACAG 240
Dy 1143 GGAGCGGAGGTCAAGGTCTACATCCAGAGCTGTGTGGAGGGCTGTACTACCTGACAG 1202
Qy 241 CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAACATCTCTGATGGTGCATCTCTCCCG 300
Dy 1203 CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAACATCTCTGATGGTGCATCTCTCCCG 1262
Qy 301 GGAAGACATTAAATCTGGGACTTTGGCTTTGGCCAGAACATCAACCCAGAGAGTGCAC 360
Dy 1263 GGAAGACATTAAATCTGGGACTTTGGCTTTGGCCAGAACATCAACCCAGAGAGTGCAC 1322
Qy 361 GTTCAGCGAGTACGGCTCCCTGAGTTCTGCTCCCGGAGATCATCCAGCAGAAACCTGT 420
Dy 1323 GTTCAGCGAGTACGGCTCCCTGAGTTCTGCTCCCGGAGATCATCCAGCAGAAACCTGT 1382
Qy 421 GAGCGAAGCCTCCGACATTTGGGCGCATGGGTGTCTCTCTACCTCAGCTGACCTGCTC 480
Dy 1383 GAGCGAAGCCTCCGACATTTGGGCGCATGGGTGTCTCTCTACCTCAGCTGACCTGCTC 1442

QY	481	ATCCCAATTGCGGCGAGAGTGAACGTCGCCACCTCTCTGTAAAGTCCTGAGAGGGGCGCGT	540
Db	1443	ATCCCAATTGCGGCGAGAGTGAACGTCGCCACCTCTCTGTAAAGTCCTGAGAGGGGCGCGT	1502
QY	541	GTCAATGAGCAGCCCATGGCTGCCACCTCAGCGAAGACGCCAAAGACTTCATCAAGGC	600
Db	1503	GTCAATGAGCAGCCCATGGCTGCCACCTCAGCGAAGACGCCAAAGACTTCATCAAGGC	1562
QY	601	TACGCTGACAGAGAGCCCTCAGGCGCGGCTAGTCGGCCAGTGCCTCTCCCAACCCCTG	660
Db	1563	TACGCTGACAGAGAGCCCTCAGGCGCGGCTAGTCGGCCAGTGCCTCTCCCAACCCCTG	1622
QY	661	GTTCTCTGAATCCATGCTCGGAGGAGGCCCATCTTCATCAACCAACCAAGCAGCTCAAGTT	720
Db	1623	GTTCTCTGAATCCATGCTCGGAGGAGGCCCATCTTCATCAACCAACCAAGCAGCTCAAGTT	1682
QY	721	CCTCTGCGCCGAAGTCGCTGCGCAGCGTTCCCTGATGAGCTACAAGTCCATCTCTGTGTAT	780
Db	1683	CCTCTGCGCCGAAGTCGCTGCGCAGCGTTCCCTGATGAGCTACAAGTCCATCTCTGTGTAT	1742
QY	781	GGGCTCCATCCCTGAGCTGCTCGGGGGCCGACCCGACAGCCCTCCTCGGCGTGAAGCCG	840
Db	1743	GGGCTCCATCCCTGAGCTGCTCGGGGGCCGACCCGACAGCCCTCCTCGGCGTGAAGCCG	1802
QY	841	GCACCTCTGCAGGGACACTGCTGGCTCCTCCAGTTCTCTCTCTCTCTGACCAACAGCT	900
Db	1803	GCACCTCTGCAGGGACACTGCTGGCTCCTCCAGTTCTCTCTCTCTCTGACCAACAGCT	1862
QY	901	CGCCCCATTTGCCCGGGTAAGTCACTGCGCAACCTCCCGGTGACACACTCAACCACTGCT	960
Db	1863	CGCCCCATTTGCCCGGGTAAGTCACTGCGCAACCTCCCGGTGACACACTCAACCACTGCT	1922
QY	961	GCACCCCCGGGCTTCCTGCGGCCCTCGGCCAGCTGCTGAGGAGCCGAGAGCCAGTGA	1020
Db	1923	GCACCCCCGGGCTTCCTGCGGCCCTCGGCCAGCTGCTGAGGAGCCGAGAGCCAGTGA	1982
QY	1021	GGGCTCAACGAGGCGCCAGCTCGCGCTGTCATCTCCGAGGGTCCGGGCCACCGGCCCG	1080
Db	1983	GGGCTCAACGAGGCGCCAGCTCGCGCTGTCATCTCCGAGGGTCCGGGCCACCGGCCCG	2042
QY	1081	CCAGGGCTGCTGTCGCCGGACACAGCTCATCCGAGGCTGTTCTACCAACAGGCGGGTGA	1140
Db	2043	CCAGGGCTGCTGTCGCCGGACACAGCTCATCCGAGGCTGTTCTACCAACAGGCGGGTGA	2102
QY	1141	GAGCCCTGAGCACGGGCGCTTGCGCCCGGGAGCAGCGGSCACCCGCGCGCGCGGCA	1200
Db	2103	GAGCCCTGAGCACGGGCGCTTGCGCCCGGGAGCAGCGGSCACCCGCGCGCGCGGCA	2162
QY	1201	CCTCTCAAGGCGGGTACATTGGGGGGCGCTGCCAGGCTTGGCGAGGCACTGATGA	1260
Db	2163	CCTCTCAAGGCGGGTACATTGGGGGGCGCTGCCAGGCTTGGCGAGGCACTGATGA	2222
QY	1261	GCACCGGCTGCTGGAGGAGGAGCGCGCCAGGAGAGCAGGCGCACCTCTCTGCCCAAGC	1320
Db	2223	GCACCGGCTGCTGGAGGAGGAGCGCGCCAGGAGAGCAGGCGCACCTCTCTGCCCAAGC	2282
QY	1321	CCCTCAATTCGAGACTGCCCTCGGCTGCTGCTGCTGGGACCCACTTGGGCCCTGGCCA	1380
Db	2283	CCCTCAATTCGAGACTGCCCTCGGCTGCTGCTGCTGGGACCCACTTGGGCCCTGGCCA	2342
QY	1381	CAGCCATCCTCTGGAACATGACTCTCCGAGCACCCCGCCGCTCTCTCGAGGCGCTCGG	1440
Db	2343	CAGCCATCCTCTGGAACATGACTCTCCGAGCACCCCGCCGCTCTCTCGAGGCGCTCGG	2402
QY	1441	TGAGGCACAGCACTGCCCTTACGCCCTTCGCGGGGGGCCCTTATCAAGGCATGGGGCA	1500
Db	2403	TGAGGCACAGCACTGCCCTTACGCCCTTCGCGGGGGGCCCTTATCAAGGCATGGGGCA	2462
QY	1501	CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGCTGCGCACCCAGGCACTGCTCAGCCAGA	1560
Db	2463	CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGCTGCGCACCCAGGCACTGCTCAGCCAGA	2522
QY	1561	GAGGCCATCCCGGACAGCCCTTGGGGGCGAGCCAGCCCTTTCTTGCCACCCCAAGCAGGG	1620

Db	2523	 GAGGCCATCCCGGACAGCCCTTGGGGGAGCCAGCCCTTCTGCAACCCCAAGCAGGG	2582
Qy	1621	TTCTGCCCCCAGGAGGGCTGCAAGCCCCCAACCAGAGTTGCCCCATGCCCCCTCTGGCTC	1680
Db	2583	TTCTGCCCCCAGGAGGGCTGCAAGCCCCCAACCAGAGTTGCCCCATGCCCCCTCTGGCTC	2642
Qy	1681	CTTCCCTCCAGGATCTTGAAAGAGGCCCCCTTAGTACCTCAAGCCCCTTCTTGGGACA	1740
Db	2643	CTTCCCTCCAGGATCTTGAAAGAGGCCCCCTTAGTACCTCAAGCCCCTTCTTGGGACA	2702
Qy	1741	GCCCAAGGCACCCCCTGCCCCCTGCCAAAGCAAGCCCCCAATTGGACTCTAAGATGGGGC	1800
Db	2703	GCCCAAGGCACCCCCTGCCCCCTGCCAAAGCAAGCCCCCAATTGGACTCTAAGATGGGGC	2762
Qy	1801	TGGAGACATCTCTCTCTGCGGAGGCCAAACCCGGCCCTGCACTCCAGAGGTCAAGC	1860
Db	2763	TGGAGACATCTCTCTCTGCGGAGGCCAAACCCGGCCCTGCACTCCAGAGGTCAAGC	2822
Qy	1861	CTCCAGGGCAGCTCTTCCCAAGTAGCTCCCTCAGGGTGGGCTCTCTCCAGGTGGGCAC	1920
Db	2823	CTCCAGGGCAGCTCTTCCCAAGTAGCTCCCTCAGGGTGGGCTCTCTCCAGGTGGGCAC	2882
Qy	1921	AGAGCTGGCCCCCTCTCTGATGCGAGGGCTGACACCAAGAGGCTGAGATCTGTCTCGA	1980
Db	2883	AGAGCTGGCCCCCTCTCTGATGCGAGGGCTGACACCAAGAGGCTGAGATCTGTCTCGA	2942
Qy	1981	CTCCACACCACTTGCAGCGGCTCAGGAACAGGCGACATGCGCAAGTTCTCCCTGGG	2040
Db	2943	CTCCACACCACTTGCAGCGGCTCAGGAACAGGCGACATGCGCAAGTTCTCCCTGGG	3002
Qy	2041	TGTTGGGGGGGCTACGCAAGGCGTGGCTGGCTATGGCACTTTGGTGGAGATGC	2100
Db	3003	TGTTGGGGGGGCTACGCAAGGCGTGGCTGGCTATGGCACTTTGGTGGAGATGC	3062
Qy	2101	AGGGGGCATGCTGGGGCAGGGGCCATGTGGGCCAGGATAGCTGGGCTGTGCCAGTC	2160
Db	3063	AGGGGGCATGCTGGGGCAGGGGCCATGTGGGCCAGGATAGCTGGGCTGTGCCAGTC	3122
Qy	2161	GGAGGAGGAGGACGAGGAGGCGCAGGGCTGAGTCCAGTGGAGGAGCAGCAGGAGGC	2220
Db	3123	GGAGGAGGAGGACGAGGAGGCGCAGGGCTGAGTCCAGTGGAGGAGCAGCAGGAGGC	3182
Qy	2221	CAGGGCTGAGAGCCACTGCCCCAGGTCAGTGCAAGGCTGTGGCTCAGTTCGGCAGGC	2280
Db	3183	CAGGGCTGAGAGCCACTGCCCCAGGTCAGTGCAAGGCTGTGGCTCAGTTCGGCAGGC	3242
Qy	2281	TCCCACCGAGGCTCTCCAGAGCCACCCCATGGGAGACA TCGGGAGGCTCTCCCTGGT	2340
Db	3243	TCCCACCGAGGCTCTCCAGAGCCACCCCATGGGAGACA TCGGGAGGCTCTCCCTGGT	3302
Qy	2341	GCAGATCCGGGACCTGTCAAGTGATGCGAGGGCGCGACACAATATCCTTGACATTC	2400
Db	3303	GCAGATCCGGGACCTGTCAAGTGATGCGAGGGCGCGACACAATATCCTTGACATTC	3362
Qy	2401	CGAGGTGACCCCGCTACCTCAACCTCTCAGACTGTACGATATCAAGTACCTCCCAT	2460
Db	3363	CGAGGTGACCCCGCTACCTCAACCTCTCAGACTGTACGATATCAAGTACCTCCCAT	3422
Qy	2461	CGAGTTTATGATCTCAGAAAGTCCCAAGTCGCTCAGCAGAGCGCCCTCCGCCAT	2520
Db	3423	CGAGTTTATGATCTCAGAAAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCGCCAT	3482
Qy	2521	GGCTGAGGAGGAGCTGGCCGAGTTCCCGAGCCCACTGGCCCTGGCAGAGCGCCCTCCGCCAT	2580
Db	3483	GGCTGAGGAGGAGCTGGCCGAGTTCCCGAGCCCACTGGCCCTGGCAGAGCGCGCTGGG	3542
Qy	2581	CCCCACCGAGGCTGGAGATCA CAGAGGAGTCAGAGGATGTGGA CGCGCTGTGGCAGA	2640
Db	3543	CCCCACCGAGGCTGGAGATCA CAGAGGAGTCAGAGGATGTGGA CGCGCTGTGGCAGA	3602
Qy	2641	GGCTCCCTGGGCAAGGCGCAAGTGGTCTCTGCGTCAAGCGCTCTTCCACTTCCC	2700

Db 725 CCGCCACGGGCTGCGTGCCTCCGGGACAGGGTCAATCCGAGCGCTGTCTTACCAACAGGCGG 784
QY 1137 GTGAGAGCCCTGAGCACGGGGCCCTGGGCCCGGGGAGCAGGCGGCAACCCGGCCGGCGGC 1196
Db 785 GTGAGAGCCCTGAGCACGGGGCCCTGGGCCCGGGGAGCAGGCGGCAACCCGGCCGGCGGC 844
QY 1197 GGCACCTGCTGAAGCGGGGTACATTTGGGGGGCGCTGCCAGGCGCTGGCGAGGCCACTGA 1256
Db 845 GGCACCTGCTGAAGCGGGGTACATTTGGGGGGCGCTGCCAGGCGCTGGCGAGGCCACTGA 904
QY 1257 TGGAGCACCGCGTGTGTGAGGAGGAGGCGCCAGGAGGAGCAGGCCACCCCTCTTGCGCA 1316
Db 905 TGGAGCACCGCGTGTGTGAGGAGGAGGCGCCAGGAGGAGGAGGCCACCCCTCTTGCGCA 964
QY 1317 AAGCCCCCTCATTTCCAGACTGCGCTCCGGCTGCGCTCTGCGCTCTGCGCACCCACTTGGCCCCCTG 1376
Db 965 AAGCCCCCTCATTTCCAGACTGCGCTCCGGCTGCGCTCTGCGCACCCACTTGGCCCCCTG 1024
QY 1377 GCCACAGCCACTCCCTGGAACATGACTCTCCGAGCACCCCGCGCCCTCTCTCGAGGCT 1436
Db 1025 GCCACAGCCACTCCCTGGAACATGACTCTCCGAGCACCCCGCGCCCTCTCTCGAGGCT 1084
QY 1437 GCGGTGAGGCACAGCGACTGCTTCAGCCCTCTCGGGGGGGCCCTCATCAGGACATGG 1496
Db 1085 GCGGTGAGGCACAGCGACTGCTTCAGCCCTCTCGGGGGGGCCCTCATCAGGACATGG 1144
QY 1497 GGCACCTCTCAGGGTCCAAAGCAGCTTCCATCCACTGTGTGGCCACCCAGGCACTCTCAGC 1556
Db 1145 GGCACCTCTCAGGGTCCAAAGCAGCTTCCATCCACTGTGTGGCCACCCAGGCACTCTCAGC 1204
QY 1557 CAGAGAGCCATCCCGGACAGCCCTTTGGGGGAGCCAGCCCTTTCTGCCACCCCAAGC 1616
Db 1205 CAGAGAGCCATCCCGGACAGCCCTTTGGGGGAGCCAGCCCTTTCTGCCACCCCAAGC 1264
QY 1617 AGGGTTCTGCCCCCAGAGGAGGCTGCAGCCGCCACCCAGCAGATTTGCCCCATGCCCTCCTG 1676
Db 1265 AGGGTTCTGCCCCCAGAGGAGGCTGCAGCCGCCACCCAGCAGATTTGCCCCATGCCCTCCTG 1324
QY 1677 GCTCTTCTCCAGGATCTTTCGAAAGAGGCCCCCTTTAGTACCCCTCAAGCCCCCTTTCTGG 1736
Db 1325 GCTCTTCTCCAGGATCTTTCGAAAGAGGCCCCCTTTAGTACCCCTCAAGCCCCCTTTCTGG 1384
QY 1737 GACAGCCCAAGCACCCTGCCCCCTGCCAAGCAGCCCAAGCCCCCAATTTGGACTCTAAGATGG 1796
Db 1385 GACAGCCCAAGCACCCTGCCCCCTGCCAAGCAGCCCAAGCCCCCAATTTGGACTCTAAGATGG 1444
QY 1797 GGCCTGGAGACATCTCTCTTCTGGGAGGCGCAAAACCCGGGCCCTGCAGTTTCCCCAGGGT 1856
Db 1445 GGCCTGGAGACATCTCTCTTCTGGGAGGCGCAAAACCCGGGCCCTGCAGTTTCCCCAGGGT 1504
QY 1857 CAGCCTCCAGGCGAGCTCTTCCCAAGTGAAGCTCCCTCAGGGTGGGCTCTCCCAAGTGG 1916
Db 1505 CAGCCTCCAGGCGAGCTCTTCCCAAGTGAAGCTCCCTCAGGGTGGGCTCTCCCAAGTGG 1564
QY 1917 GCACAGAGCTGGCCCCCTCTCTGATGCGAGGAGGCTGGACCCAGGAGGCTGAGGATCTGT 1976
Db 1565 GCACAGAGCTGGCCCCCTCTCTGATGCGAGGAGGCTGGACCCAGGAGGCTGAGGATCTGT 1624
QY 1977 CCGACTCCACACCCACTTTCAGCGGCTCAGGAAACAGGCGACCATGCGCAAGTTCTCCC 2036
Db 1625 CCGACTCCACACCCACTTTCAGCGGCTCAGGAAACAGGCGACCATGCGCAAGTTCTCCC 1684
QY 2037 TGGGTGCTCGGGGGGTACGCAAGGCTGCTGGCTATGCGACCTTTGCTTTGGTGGAG 2096
Db 1685 TGGGTGCTCGGGGGGTACGCAAGGCTGCTGGCTATGCGACCTTTGCTTTGGTGGAG 1744
QY 2097 ATCAGGGGGCATGCTGGGCGAGGGGCCATGTGGGCCAGGATAGCTGGGCTGTCTCC 2156
Db 1745 ATCAGGGGGCATGCTGGGCGAGGGGCCATGTGGGCCAGGATAGCTGGGCTGTGTCTCC 1804
QY 2157 AGTCGAGGAGGAGGAGCAGGAGGAGCCAGGCTGAGTCCCAAGTCCGAGGAGCAGCAGG 2216
Db 1805 AGTCGAGGAGGAGGAGCAGGAGGAGCCAGGCTGAGTCCCAAGTCCGAGGAGCAGCAGG 1864

QY 2217 AGGCCAGGCTGAGAGGCCCACTGCCCCCAGGTCAGTGCAAGGCCCTGTGCTGAGGTCGSCA 2276
Db 1865 AGGCCAGGCTGAGAGGCCCACTGCCCCCAGGTCAGTGCAAGGCCCTGTGCTGAGGTCGSCA 1924
QY 2277 GGGTCTCCACAGAGGCTCTCCAGAGCCCAACCCCATCGGAGGACATCGGAGAGTCTCCC 2336
Db 1925 GGGTCTCCACAGAGGCTCTCCAGAGCCCAACCCCATCGGAGGACATCGGAGAGTCTCCC 1984
QY 2337 TGGTGAGATCCGGGACCTGTCAAGTGTATCGGAGGCGCGGACACAATATCCCTGAGCA 2396
Db 1985 TGGTGAGATCCGGGACCTGTCAAGTGTATCGGAGGCGCGGACACAATATCCCTGAGCA 2044
QY 2397 TTTCCGAGGTGGACCCCGCCCTACTCAACTCTCAGACCTGTACGATATCAAGTACCTCC 2456
Db 2045 TTTCCGAGGTGGACCCCGCCCTACTCAACTCTCAGACCTGTACGATATCAAGTACCTCC 2104
QY 2457 CATTCGAGTTTATGATCTTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCCGCCCTCCC 2516
Db 2105 CATTCGAGTTTATGATCTTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCCGCCCTCCC 2164
QY 2517 CCATGCTGAGGAGGAGCTGGCCGAGTTCCCGAGGCCCAAGTGGCCCTGGCCAGGTAAC 2576
Db 2165 CCATGCTGAGGAGGAGCTGGCCGAGTTCCCGAGGCCCAAGTGGCCCTGGCCAGGTAAC 2224
QY 2577 TGGGCCCCCAACGAGGCTGGAGATCACAGAGGAGTCAGAGGATGTGGACGCGCTGCTGG 2636
Db 2225 TGGGCCCCCAACGAGGCTGGAGATCACAGAGGAGTCAGAGGATGTGGACGCGCTGCTGG 2284
QY 2637 CAGAGCTGCGTGGGCGAGGAGCGCAAGTGTCTTCCGCTCAGCGAGCTCTTCCACT 2696
Db 2285 CAGAGCTGCGTGGGCGAGGAGCGCAAGTGTCTTCCGCTCAGCGAGCTCTTCCACT 2344
QY 2697 TCCTCGGAGGACCTGCGCTGGATGAGCTGAGGCTGGGCTGGGCTGGGCTGGGCTGGG 2756
Db 2345 TCCTCGGAGGACCTGCGCTGGATGAGCTGAGGCTGGGCTGGGCTGGGCTGGGCTGGG 2404
QY 2757 AGSCCTCCGTGGAGCACATCTCCCGGATCTTGAAGGCGAGGCGGGAAGGTCTGGAGAAG 2816
Db 2405 AGSCCTCCGTGGAGCACATCTCCCGGATCTTGAAGGCGAGGCGGGAAGGTCTGGAGAAG 2464
QY 2817 AGSGGCCCCCAGGAAAGACAGGCTTCTTCCGCTCTCAGGCTCTCAGGCTTGAAGCT 2876
Db 2465 AGSGGCCCCCAGGAAAGACAGGCTTCTTCCGCTCTCAGGCTCTCAGGCTTGAAGCT 2524
QY 2877 GGGACGAGCGCGGACATTCCTAAGGAGCTCTCAGATCAGACTGTGTCTGGGCGAGT 2936
Db 2525 GGGACGAGCGCGGACATTCCTAAGGAGCTCTCAGATCAGACTGTGTCTGGGCGAGT 2584
QY 2937 CAGTGACACTGCGCTGCCAGGTGTCAAGCCAGCCAGCTGCCAGGCGACCTGGAGCAAG 2996
Db 2585 CAGTGACACTGCGCTGCCAGGTGTCAAGCCAGCCAGCTGCCAGGCGACCTGGAGCAAG 2644
QY 2997 ACGGA 3001
Db 2645 ACGGA 2649

RESULT 10

AAD26467
ID AAD26467 standard; cDNA; 3225 BP.

XX AAD26467;

AC AAD26467;

XX AAD26467;

DT 26-MAR-2002 (first entry)

XX Human kinase PKIN-20 cDNA.

DE Human; kinase; PKIN-20; cancer; leukaemia; adenocarcinoma; osteoporosis;

XX Human; kinase; PKIN-20; cancer; leukaemia; adenocarcinoma; osteoporosis;

KW immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;

KW Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;

KW allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;

KW autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;

QY 2360 GGTGATGCGAGGCGCGGACACAATATCCCTTGGACATTTCCGAGGTGGACCCCGCCTAC 2419
DB |||||
QY 1093 GGTGATGCGAGGCGCGGACACAATATCCCTTGGACATTTCCGAGGTGGACCCCGCCTAC 1152
DB |||||
QY 2420 CTCACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTCGAGTTTATGATCTTACG 2479
DB |||||
QY 1153 CTCACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTCGAGTTTATGATCTTACG 1212
DB |||||
QY 2480 AAGTCCCAAGTCCGCTCAGCCAGAGCGCGCCTCCCCCATGGCTGAGGAGGAGCTGSCC 2539
DB |||||
QY 1213 AAGTCCCAAGTCCGCTCAGCCAGAGCGCGCCTCCCCCATGGCTGAGGAGGAGCTGSCC 1272
DB |||||
QY 2540 GAGTTCCCGAGGCCACCTGCGCCTTGGCCAGGTGAATCTGGCCGCCACCGAGGCTTGGAG 2599
DB |||||
QY 1273 GAGTTCCCGAGGCCACCTGCGCCTTGGCCAGGTGAATCTGGCCGCCACCGAGGCTTGGAG 1332
DB |||||
QY 2600 ATCAGAGGAGTCAAGAGGATGAGCGCGCTGTGGCAGAGGCTGCCGTGGGAGGAG 2659
DB |||||
QY 1333 ATCAGAGGAGTCAAGAGGATGAGCGCGCTGTGGCAGAGGCTGCCGTGGGAGGAG 1392
DB |||||
QY 2660 CGCAAGTGGTCTCGCCCTCAGCAGAGCTCTTCCACTTCCCTGGGAGGACCTGCCGCTG 2719
DB |||||
QY 1393 CGCAAGTGGTCTCGCCCTCAGCAGAGCTCTTCCACTTCCCTGGGAGGACCTGCCGCTG 1452
DB |||||
QY 2720 GATGAGCTTCAGAGCTGGGCTGCGTGAGAGAGTGAAAGGCTCCGTGGAGCACATCTCC 2779
DB |||||
QY 1453 GACGAGCTTCAGAGCTGGGCTGCGTGAGAGAGTGAAAGGCTCCGTGGAGCACATCTCC 1512
DB |||||
QY 2780 CGGATCTCTGAAGGCGAGCGCGGAGGCTCTGAGAGGAGGCGGCGGCGGAGGAGGCGCA 2839
DB |||||
QY 1513 CGGATCTCTGAAGGCGAGCGCGGAGGCTCTGAGAGGAGGCGGCGGCGGAGGAGGCGCA 1572
DB |||||
QY 2840 GGCCTTGTCTCTTCCGCTCTCAGGCTCTGAGGCTGGGAGCGGAGCGGAGGAGGAGGAGGAG 2899
DB |||||
QY 1573 GGCCTTGTCTCTTCCGCTCTCAGGCTCTGAGGCTGGGAGCGGAGGAGGAGGAGGAGGAG 1632
DB |||||
QY 2900 AGGAGCTCTCAGATGAGACTGTGGTCTGGGCCAGTCACTGAGCACTGGGCTGCCAGGTG 2959
DB |||||
QY 1633 AGGAGCTCTCAGATGAGACTGTGGTCTGGGCCAGTCACTGAGCACTGGGCTGCCAGGTG 1692
DB |||||
QY 2960 TCAGCCAGCCAGCTGCCAGGCCACCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3001
DB |||||
QY 1693 TCAGCCAGCCAGCTGCCAGGCCACCTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1734
DB |||||
RESULT 11
ID AAF44662
XX AAF44662 standard; cdna; 7710 BP.
XX
XX AAF44662;
AC
XX
XX 27-MAR-2001 (first entry)
DT
DE Novel protein kinase cdna, SEQ ID NO: 42.
XX
XX Human; mouse; protein kinase; antiarthritic; antisclerotic; osteopathic;
KW immunosuppressive; cardiant; renal; antiinflammatory; antiasthmatic;
KW dermatological; antidiabetic; antiinfertility; gene therapy; vaccine;
KW immune disorder; cardiovascular disease; neurodegenerative disease;
KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200073469-A2.
PN
XX
XX 07-DEC-2000.
PD
XX
XX 26-MAY-2000; 2000WO-US014842.
PF
XX
XX 28-MAY-1999; 99US-0136503P.
PR
XX
XX

PA
XX
PI
XX
XX
DR
XX
XX
XX
PT
XX
XX
XX
PS
XX
XX
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
CC
XX
XX

(SUGE-) SUGEN INC.
Plowman GD, Martinez R, Whyte D, Sudersanam S;
WPI; 2001-032161/04.
P-PSDB; AAB65635.

Nucleic acids encoding kinase polypeptides, useful for diagnosing and treating immune-related diseases and disorders, cardiovascular disease, neurodegenerative diseases and/or cancers.

Example 1; Fig 2; 310pp; English.

The present sequence encodes a novel protein kinase. The nucleic acids and the protein kinases they encode may be used in the treatment and diagnosis of diseases associated with inappropriate kinase expression such as immune-related diseases and disorders, cardiovascular disease, neurodegenerative diseases and/or cancers. The nucleic acids and complementary sequences may also be used as DNA probes in diagnostic assays. The kinase polypeptides may be used as antigens in the production of antibodies of kinase expression and activity. Anti-kinase antibodies of kinases antagonists may also be used to down regulate kinase expression and activity. Diseases related to kinase expression and activity include rheumatoid arthritis, atherosclerosis, autoimmune disorders, complications of organ transplantation, myocardial infarction, immune disorders, cardiomyopathies, strokes, renal failure, oxidative-stress related disorders, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive disorders

Sequence 7710 BP; 1446 A; 2534 C; 2447 G; 1283 T; 0 U; 0 Other;

Query Match 7.4%; Score 223.2; DB 4; Length 7710;
Best Local Similarity 59.0%; Pred. No. 1.2e-30;
Matches 466; Conservative 0; Mismatches 303; Indels 21; Gaps 4;
QY 33 GGGAGCGAGACATCTCTGGCGGCTGAGCCACCGCTGGTCAAGGCTGCTGGACAGT 92
DB 1988 GGGAGCGCGGCTGCTGGCGAGCTCCAGCAGACGTGTCTCTACTTCCATGAGGCCT 2047

QY 93 TTGAGACCCGCAAGACCCCTCATCTCTGAGAGCTGTGCTCATCCGAGGAGCTGCTGG 152
DB 2048 TCGAGAGGCGCGGGAGCTGGTCAATTGTCAACGAGCTCTGCACA---GAGGAGCTGCTGG 2104

QY 153 ACCGCTGTACAGGAAGGCGGTGTGACGAGGCGAGGTCAGAGTCTACATCCAGCAGC 212
DB 2105 AGCGAATCCCAAGAAACCCACCGTGTGTGATCTGAGATCCGGGCTATATGCGGCGAG 2164

QY 213 TGGTGGAGGGCTGCACATACTGTGCACAGCCATGGCGTTCTCCACCTGGACATAAAGCCCT 272
DB 2165 TGCTAGAGGGAATACACTACCTGCACAGAGCCAGTGTGCACCTCGATGTCAAGCCTG 2224

QY 273 CTAAACATCTGATGTGCATCTCTGC-----CGGGAAGACATTAATCTGGACTTTG 326
DB 2225 AGAACCTGTGTGTGGGATGGTGTGCGGGGAGCAGCAGGTGCGGATCTGTGACTTTG 2284

QY 327 GCTTTGCCAGAAACATCACCCAGCAGAGCTCAGTTACAGCCAGTACGGCTCCCTGAGT 386
DB 2285 GGAATGCCGAGGAGTGAATCCAGAGAGCCCGAGTACTGCCAGTATGTCACACCTGAGT 2344

QY 387 TCGTCTCCCGGAGATCATCCAGCAGAAACCTGTGTGAGCGAAGCCTCCGACATTTGGGCCA 446
DB 2345 TTGTAGCACCCGAGATTGTCAATCAGAGCCCGTGTCTGGAGTCACTGACATCTGSCCTG 2404

QY 447 TGGGTGTCTCTCTACCTCAGCTGTCTCATCCCCATTTGCCGCGAGAGTGACC 506
DB 2405 TGGGTGTGTGCTCTCTCTGTCTGTGACAGGAATCTCCCCGTTTGTGGGAAAAATGACC 2464

QY 507 GTGCCACCCCTCTGTAAACGTCCTGGAGGGGCGGTGTGATGGAGCAGGCCCATGGTGCCC 566
DB 2465 GGACAAACATTGATGAACATCCGAACTACAGCTGGCCTTCGAGGAGACCAATTCCTGA 2524

XX Human; kinase polypeptide; PKIN-18; gene therapy; Addison's disease;
 KW leukemia; immune disorder; lymphoma; melanoma; developmental disorder;
 KW asthma; Crohn's disease; rheumatoid arthritis; bursitis; atherosclerosis;
 KW cirrhosis; hepatitis; psoriasis; Cushing's syndrome; cytostatic; cancer;
 KW cholestasis; cardiant; cardiovascular disorder; Niemann-Pick's disease;
 KW lipid disorder; fatty liver; Gaucher's disease; myocardial infarction;
 KW drug screening; transgenic animal; antiinflammatory; hepatotropic;
 KW hypotensive; anti-HIV; enzyme; ss.
 XX
 OS Homo sapiens.
 XX
 XX WO200208399-A2.
 XX
 XX 31-JAN-2002.
 XX
 XX 20-JUL-2001; 2001WO-US023092.
 XX
 XX 21-JUL-2000; 2000US-0220038P.
 PR 28-JUL-2000; 2000US-0222112P.
 PR 04-AUG-2000; 2000US-0222831P.
 PR 11-AUG-2000; 2000US-0224729P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA (THOR/) THORNTON H.
 XX
 XX Thornton M, Yue H, Khan FA, Gururajan R, Hafalia AJA, Wallia NK;
 PI Patterson C, Ramkumar J, Gandhi AR, Policky JL, Baughn MR;
 PI Tribouley CM, Bandman O, Nguyen DB, Lu Y, Burford N, Lal P, Ding L;
 PI Yao MG, Elliott VS, Recipon SA, Kearney L, Lu DAM, Greenwald SR;
 PI Tang YT, Xu Y, Walsh RT, Gietzen KJ, Yang J, Hillman JL;
 XX WPI; 2002-206083/26.
 DR
 XX
 XX New human kinase polypeptide, useful in diagnosis, prevention and
 PT treatment of cancer, immune disorder, growth and developmental disorder,
 PT cardiovascular disorder and lipid disorder.
 XX
 XX Claim 5; Page 191-193; 196pp; English.
 XX
 XX The present invention relates to an isolated human kinase polypeptide
 CC (PKIN) or a biologically active/immunogenic fragment of PKIN. PKIN is
 CC useful for diagnosing, treating and preventing cancer (e.g., leukaemia,
 CC lymphoma, melanoma), an immune disorder (e.g., acquired immunodeficiency
 CC syndrome (AIDS), Addison's disease, allergy, anaemia, asthma, Crohn's
 CC disease, rheumatoid arthritis), a growth and developmental disorder (e.g.
 CC bursitis, cirrhosis, hepatitis, psoriasis, Cushing's syndrome), a
 CC cardiovascular disorder (e.g., atherosclerosis, hypertension, myocardial
 CC infarction), and a lipid disorder (e.g., fatty liver, cholestasis,
 CC Gaucher's disease, Niemann-Pick's disease). PKIN is useful in a number of
 CC drug screening techniques and to analyse the proteome of a tissue or cell
 CC type. PKIN is useful for creating knockin humanised animals or transgenic
 CC animals to model human diseases, in somatic or germline gene therapy, to
 CC generate a transcript image of a tissue or cell type, for detecting
 CC differences in the chromosomal location due to translocation, inversion,
 CC etc., among normal, carrier or affected individuals, and as hybridisation
 CC probes for mapping naturally occurring genomic sequences. PKIN is useful
 CC in southern or northern analysis, dot blot or other membrane-based
 CC technologies, in PCR technologies, in dipstick, pin, microarray enzyme
 CC linked immunosorbent (ELISA)-like assays and in microarrays utilising
 CC fluids or tissues from patients to detect altered PKIN expression. The
 CC present sequence is human PKIN-18 cDNA. Note: This sequence is said to
 CC encode PKIN-18 referred as SEQ ID NO:18 (AAE19160). However this does not
 CC appear to be the case
 XX
 XX Sequence 7789 BP; 1469 A; 2558 C; 2470 G; 1292 T; 0 U; 0 Other;
 SQ
 Query Match 7.4%; Score 223.2; DB 6; Length 7789;
 Best Local Similarity 59.0%; Pred. No. 1.2e-30;
 Matches 466; Conservative 0; Mismatches 303; Indels 21; Gaps 4;
 33 GGGAGCGAGACATCTCTGGCGCGGTGAGCCACCCGCTGTGTCACGGGGCTGTGGAGCAGT 92

Db	2050	GGGAGGCGCGGCTGCTGGCAGGGTCCAGCAGCACTGTGCTCTACTTCCATGAGGCCT	2109
Qy	93	TTGAGACCCGCAAGACCCCTCATCTCTCTGGAGCTGTGCTCATCCGAGGAGCTGCTGG	152
Db	2110	TCGAGAGCGCGGGGACTGGTCAATTGTACCCGAGCTCTGCACA---GAGAGCTGCTGG	2166
Qy	153	ACGCGCTGTACAGAAAGGGGCTGGTGTAGCGAGGCGGAGGTCAAGGTCTACATCCAGCAGC	212
Db	2167	AGCGAATCGCCAGGAAACCCACCGTGTGTGAGTCTGAGATCCGGGCGCTATATCGGCGAGG	2226
Qy	213	TGCTGAGAGGGCTGCACCTACCTGCACAGCCATGCGGCTTCTCCACCTGGACATAAAGCCCT	272
Db	2227	TGCTAGAGGGAATACACTACTGTGACACAGACCACTGTGTGCACCTGTGATGTCAAGCTG	2286
Qy	273	CTAACATCTGTGATGGTGCATCTCTGC-----CCGGGAAGACATTAATAATTCGGACTTTG	326
Db	2287	AGAACCTGTGTGTGGATGGTGTGCGGGCGAGCAGAGGTGCGGATCTGTGACTTTG	2346
Qy	327	GCTTTGCCAGAACATCACCCAGCAGAGCTGCAGTTTCAGCCAGTAGGGTCCCTTGAGT	386
Db	2347	GGATGCCAGGAGCTGACTCCAGGAGAGCCCGACTGTGCCAGTATGGCACACTGAGT	2406
Qy	387	TCGTCTCCCCGAGATCATCCAGCAAAACCTGTGTAGCGAAGCCTCCGACATTTGGGCCA	446
Db	2407	TTGTAGCACCCGAGATTGTCAATCAGAGCCCGTGTCTGGAGTCACTGACATCTGSCCTG	2466
Qy	447	TGGGTGTCTCTCTACCTCAGCTGACCTGTCTATCCCAATTTGCCGGGAGAGTGACC	506
Db	2467	TGGGTGTGTGTGCTTCTCTCTGTGTACAGGAATCTCCCGCTTTGTTGGGAAAAATGACC	2526
Qy	507	GTGCCACCTCTCTGAACGTCTTGGAGGGCGGTGTCTGAGCGAGAGCCCGATGGCTGCC	566
Db	2527	GGACACATTTGATGAACATCCGAACTCAACGTGGCCTTCGAGGAGACCAATTCCTGA	2586
Qy	567	ACCTCAGCAAGACGCCAAAGACTTTCATCAAGGCTACGCTGCAGAGAGCCCTCAGGCC	626
Db	2587	GCTCAGCAGGAGGCGCGGGCTTCTCATCAAAGTGTGTGTGCGAGGACCGGCTG---A	2643
Qy	627	GGCTAGTGGCGCCAGTGCCCTCTCCACCCCTGGTCTCTGAAATCCATGCTGGGAGG	686
Db	2644	GACCTACCGCAAGAGACCCCTAGAACATCTCTTGGT-----CAAAATCAGGCNA	2694
Qy	687	AGGCCCCATTTCAATCAACACCAAGCAGCTCAAGTTCCTCTGCGCCGAAAGTCTGCGCAGC	746
Db	2695	AGGGCGCAGAGGTGAGCAGCGGATCACTGAAGCTATTCTCTCCCGCGGAGGTGGCAGC	2754
Qy	747	GTTCCTGTATGAGTACAAAGTCCATCTCTGGTGTATGGCTTCCATCCCTGAGCTGTGCGGG	806
Db	2755	GCTCCCGATCAGCTACAAATGCCCACCTGGTGTGCTGCGGCCCATCCCGAGCTGTGCGGG	2814
Qy	807	GCCCCACCGCA 816	
Db	2815	CCCCCCCAGA 2824	
RESULT 14			
ADE47675			
ID	ADE47675 standard; DNA; 9698 BP.		
XX	ADE47675;		
AC	ADE47675;		
XX	29-JAN-2004 (first entry)		
DT	XX		
XX	Human NOV14c gene SEQ ID NO:37.		
DE	XX		
KW	ds; gene; human; cardiant; antiarteriosclerotic; hypotensive;		
KW	immunosuppressive; dermatological; anorectic; cytostatic; antidiabetic;		
KW	haemostatic; anti-HIV; antiasthmatic; antibacterial; virucide;		
KW	neuroprotective; nootropic; antiparkinsonian; antilipaeamic; gene therapy;		
XX	vaccine.		
OS	Homo sapiens.		

XX PN WO2003076642-A2.

XX PD 18-SEP-2003.

XX PF 02-AUG-2002; 2002WO-US024459.

XX PR 02-AUG-2001; 2001US-0309501P.

XX PR 03-AUG-2001; 2001US-0310291P.

XX PR 08-AUG-2001; 2001US-0310951P.

XX PR 09-AUG-2001; 2001US-0311292P.

XX PR 13-AUG-2001; 2001US-0311979P.

XX PR 14-AUG-2001; 2001US-0312203P.

XX PR 17-AUG-2001; 2001US-0313156P.

XX PR 17-AUG-2001; 2001US-0313201P.

XX PR 20-AUG-2001; 2001US-0313702P.

XX PR 21-AUG-2001; 2001US-0314031P.

XX PR 23-AUG-2001; 2001US-0314466P.

XX PR 28-AUG-2001; 2001US-0315403P.

XX PR 29-AUG-2001; 2001US-0315853P.

XX PR 31-AUG-2001; 2001US-0316508P.

XX PR 21-SEP-2001; 2001US-0323936P.

XX PR 03-DEC-2001; 2001US-0338078P.

XX PR 05-FEB-2002; 2002US-0334655P.

XX PR 05-MAR-2002; 2002US-0361764P.

XX PR 19-APR-2002; 2002US-0373825P.

XX PR 15-MAY-2002; 2002US-0380971P.

XX PR 15-MAY-2002; 2002US-0380980P.

XX PR 16-MAY-2002; 2002US-0381039P.

XX PR 28-MAY-2002; 2002US-0383761P.

XX PR 29-MAY-2002; 2002US-0383887P.

XX PR 01-AUG-2002; 2002US-00210130.

XX PA (CURA-) CURAGEN CORP.

XX PI Zehrusen BD, Patturajan M, Kekuda R, Miller CE, Rieger DK;

XX PI Pena CE, Shimkets RA, Li L, Berghs C, Zhong M, Caeman SJ, Voss EZ;

XX PI Boldog FL, Padigaru M, Smithson G, Shenoy SG, Ji W, Gorman L;

XX PI Vernet CAM, Leite MW, Guo X, Anderson DW, Spytek KA, Gerlach VL;

XX PI Burgess CB, Khrantsov NV, Ort T, Ellerman K, Rastelli L, Agee ML;

XX PI Chaudhuri A, Chant JS, Dipippo VA, Edinger SR, Eisen A, Gangolli EA;

XX PI Giot L, Ooi CE, Rothenberg ME, Spaderna SK, Hjalt T, Liu X;

XX PI Taupier RJ, Catterton E;

XX PI WPI; 2003-779062/73.

XX DR P-PSDB; ADE47676.

XX PT New NOVX polypeptides and nucleic acids, useful for preventing or

XX PT treating NOVX-associated disorders, e.g. cancer, diabetes,

XX PT atherosclerosis, asthma or AIDS, and in chromosome mapping, tissue typing

XX PT or pharmacogenomics.

XX PS Claim 20; SEQ ID NO 37; 562pp; English.

XX CC The invention relates to a novel (NOVX) human polypeptide. A polypeptide

XX CC of the invention has cardiant, antiarteriosclerotic, hypotensive,

XX CC immunosuppressive, dermatological, anorectic, cytostatic, antidiabetic,

XX CC haemostatic, anti-HIV, antiasthmatic, antibacterial, virucide,

XX CC neuroprotective, nontropic, antiparkinsonian, and antilipaemic activity.

XX CC A polynucleotide encoding a polypeptide of the invention may have a use

XX CC in gene therapy, and as a vaccine. A polypeptide of the invention is

XX CC useful in the manufacture of a medicament for treating a syndrome

XX CC associated with a human disease, the disease selected from a pathology

XX CC associated with the polypeptide. These may also be used in diagnosing,

XX CC treating or preventing NOVX-associated disorders such as cardiomyopathy,

XX CC atherosclerosis, hypertension, scleroderma, obesity, cancer, diabetes,

XX CC haemophilia, graft-versus-host disease, AIDS, asthma, Crohn's disease,

XX CC multiple sclerosis, infections, anorexia, cancer-associated cachexia,

XX CC neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's

XX CC disease), haematopoietic disorders, dyslipidaemias and other wasting

XX CC disorders associated with chronic diseases. The nucleic acids are also

XX CC used as hybridisation probes, in chromosome mapping, tissue typing,

XX CC preventive medicine, and pharmacogenomics. The polypeptides are also

CC useful as vaccines. The present sequence encodes a NOVX polypeptide of

CC the invention.

CC Sequence 9698 BP; 1693 A; 3352 C; 3163 G; 1490 T; 0 U; 0 Other;

XX SQ

Query Match 7.4%; Score 223.2; DB 10; Length 9698;

Best Local Similarity 59.0%; Pred. No. 1.2e-30;

Matches 466; Conservative 0; Mismatches 303; Indels 21; Gaps 4;

QY 33 GGGAGCGAGACATCTCTGGCCGCTGAGCCACCGCTGGTTCACGGGGCTGTGGACCAAGT 92

DB 4814 GGGAGGCCGCTCTGGCCAGGCTCCAGCAGCACTGTGTCTTCTTCCATGAGGCT 4873

QY 93 TTGAGACCCGCAAGACCTCATCTCTCTGGAGCTGTCTCATCCGAGGAGCTGCTGG 152

DB 4874 TCGAGAGCGCGCGGAGCTGGTCAATGTGTACCGAGCTCTGCACA---GAGGAGCTGCTGG 4930

QY 153 ACCGCTGTACAGGAAGGGCTGTGTACGAGGCGGAGGTCAAGTCTTACATCCAGCAGC 212

DB 4931 AGCGAATCGCCAGGAAACCCACCGTGTGTAGTCTGAGATCCGGGCTATATCGGCAGG 4990

QY 213 TGGTGGAGGGCTGCATCTACCTGCACAGCCATGGCTTTCACCTGGACATAAAGCCCT 272

DB 4991 TGCTAGAGGGAATACACTACCTGCACAGACCCAGCTGCTGCACCTCGATGCAAGCCTG 5050

QY 273 CTAACATCTCTGATGGTGCATCTCTG-----CCGGGAAGACATTAAATCTGCGACTTTG 326

DB 5051 AGAACCTGTGTGTGGGATGGTCTCGGGCGAGCAGAGTGGGATCTGTGACTTTG 5110

QY 327 GCTTTGCCAGAACATCACCCACAGAGCTGTCAGTTTCAGCCAGTACGGCTCCCTCGAGT 386

DB 5111 GGAATGCCAGGAGCTGACTCCAGAGAGCCCGGAGTACTGCGCAGTATGGCACCTGAGT 5170

QY 387 TCGTCTCCCCGAGATCATCCAGAGAACCTGTGTAGCGAAGCTTCGACATTTGGGGCA 446

DB 5171 TTGTAGCACCCGAGATTGTCAATCAGAGCCCGTGTCTGGAGTCACTGACATCTGGCCTG 5230

QY 447 TGGTGTCACTCTCTACTGACCTGACCTGCTCATCCCCATTTCCCGCGGAGAGTGACC 506

DB 5231 TGGGTGTGTGTGGCTTCTCTGTCTGACAGGAATCTCCCCGTTTGTGGGAAAATGACC 5290

QY 507 GTGCACCCCTCTCAAGCTCTCTGGAGGGCGCTGTGTATGAGCAGCCCAATGGCTGCC 566

DB 5291 GGACAAATTGATGAACATCCGAACTTCAACGTGGCTTCGAGAGAGCCACATTCCTGA 5350

QY 567 ACCTCAGCGAAGACCGCAAGAGCTTCAAGGCTACGCTGCAGAGAGCCCTCAGGCC 626

DB 5351 GCCTGAGCAGGAGGCGCGGGCTTCTCTCATCAAGTGTGGTGCAGGACCGCTG---A 5407

QY 627 GGCTAGTGGCGCCAGTGGCTCTCCACCCCTGGTTCCTGAAATCCATGCTGGGAGG 686

DB 5408 GACCTACCGCAGAGAGACCTTAGAACATCTTGGTT-----CAAAACTCAGGCAG 5458

QY 687 AGGCCCACTTCATCAACCAAGCAGCTCAAGTTCCTCTGGCGCGAAGTCTGCTGGCAGC 746

DB 5459 AGGGCGCAGAGGTGAGCAGCGGATCACTGAAGCTATTCTCTCCGGCGAGGTGGCAGC 5518

QY 747 GTTCCCTGATGAGCTACAAGTCCATCTCTGGTGTGCGCTCCATCCCTGAGCTGCTGGG 806

DB 5519 GCTCCAGATCAGCTACAAATGCCACCTGGTGTGCTGGCGCCCATCCCCGAGCTGCTGGG 5578

QY 807 GCCCACCCGA 816

DB 5579 CCCCCCAGAG 5588

RESULT 15

ADJ78945

ID ADJ78945 standard; DNA; 9698 BP.

XX AC ADJ78945;

XX DT 06-MAY-2004 (first entry)

PA	(AGEE//)	AGEE M L.
PA	(CHAU//)	CHAUDHURI A.
PA	(CHAN//)	CHANT J S.
PA	(DIEI//)	DIPPIO V A.
PA	(EDIN//)	EDINGER S R.
PA	(EISE//)	EISEN A J.
PA	(GANG//)	GANGOLLI E A.
PA	(GIOT//)	GIOT L.
PA	(OOIC//)	OOI C E.
PA	(ROTH//)	ROTHENBERG M E.
PA	(SPAD//)	SPADERNA S K.
PA	(HUAL//)	HUALT T.
PA	(LUX//)	LIU X.
PA	(TAUP//)	TAUPIER R J.
PA	(CATT//)	CATTERTON E.
PA	(SHEN//)	SHENOY S G.
XX		
PI	Zerhusen BD,	Patturajan M, Kekuda R, Miller CE, Rieger DK;
PI	Bena CBA,	Shimkets RA, Li L, Berghs C, Zhong M, Casman SJ, Voess EZ;
PI	Boldog FL,	Padigar M, Smithson G, Ji W, Gorman L, Vernet CAM;
PI	Leite MW,	Guo XS, Anderson DW, Spytek KA, Gerlach V, Burgess CE;
PI	Khamteov NV,	Ort T, Ellerman K, Rastelli L, Acee ML, Chaudhuri A;
PI	Chant JS,	Dipippo VA, Edinger SR, Eisen AJ, Gangolli EA, Glot L;
PI	Ooi CE,	Rotherberg ME, Spaderna SK, Hjalt T, Liu X, Taupier RJ;
PI	Catterton E,	Shenoys SG;
XX		
DR	WPI; 2004-108206/11.	
DR	P-PSDB; ADJ78946.	
XX		
PT	New isolated NOVX polypeptides and nucleic acid molecules useful for	
PT	treating, preventing and diagnosing pathological conditions with NOVX-	
PT	associated disorders, such as cancer, obesity, diabetes and inflammatory	
PT	or CNS diseases.	
XX		
PS	Claim 20; SEQ ID NO 37; 250pp; English.	
XX		
CC	This invention relates to a novel isolated NOVX polypeptide comprising a	
CC	fully defined sequence of, a mature form, one or more conservative	
CC	substitutions or at least 95% identity to 247 amino acids as given in the	
CC	specification. The invention may be useful for the development of	
CC	compounds with a cytostatic, antidiabetic, anorectic, cerebroprotective,	
CC	neuroprotective, antiinflammatory, thyromimetic or cardiant activity. In	
CC	addition, the disclosed sequences may prove useful for gene-therapy or	
CC	antisense-therapy. The invention may be useful for the diagnosis and	
CC	treatment of disorders associated with aberrant expression or activity of	
CC	the NOVX polypeptide, such as cancer, diabetes, obesity, and endocrine,	
CC	CNS, cardiovascular and inflammatory disorders. They can also be used in	
CC	various detection and screening assays, chromosome mapping, tissue typing	
CC	and predictive medicine. The present sequence is that of a gene which	
CC	encodes a human NOVX protein of the invention.	
XX		
SQ	Sequence 9698 BP; 1693 A; 3352 C; 3163 G; 1490 T; 0 U; 0 Other;	
	Query Match 7.4%; Score 223.2; DB 12; Length 9698;	
	Best Local Similarity 59.0%; Pred. No. 1.2e-30;	
	Matches 466; Conservative 0; Mismatches 303; Indels 21; Gaps 4;	
QY	33 GGGAGCGAGACATCTTCGCCGCGGTGAGACCACCCCGTGTTGCACGGGGCGCTGTGCACAGT 92	
Db	4814 GGGAGCCCGCGGTGTGTGCCAGGGTCCAGCACGACTGTGTCTTCTACTTCATGAGCGCT 4873	
QY	93 TTGAGACCCGCAAGACCTCATCTCTCATCTCTGGAGCTGTGCTCATCCGAGGAGCTGCTGG 152	
Db	4874 TCGAGAGGCGCGGGAGCTGTGTCATTTGTCACCGAGCTCTTGACA--GAGAGCTGCTGG 4930	
QY	153 ACCGCTGTGACGAAGGGCGGTGTGAGCGGAGGCCGAGGTTCAAGGTTCATATCCAGCAGC 212	
Db	4931 AGCGAATCGCCAGGAAACCCACCGTGTGTGAGTCTGAGATCCGGGCCCTATATCGGCAGG 4990	
QY	213 TGSTGGAGGGGTGCACCTACCTCTGCACGCCATGGCGTTCTCCACCTGCATTAAGCCT 272	
Db	4991 TCGTAGGGGAATACACTTACTCTGCACGAGCCACGTGCTGCACCTCGATGTCAAGCGTG 5050	


```
Qy 273 CTAACATCCTGATGGTGCATCCTGC-----CCGGAGAGACATTAAATCTGCGACTTTG 326
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5051 AGAACCTGCTGCTGTGGGATGGTGTGCGGCGAGCAGCAGGTGCGGATCTGTGACTTTG 5110
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 327 GCTTTGCCAGAAACATCACCCAGCAGAGCTGCAAGTTTCAGCCAGTACGGCTCCCTGAGT 386
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5111 GGAATGCCAGGAGCTGACTCCAGGAGAGCCCCAGTACTGCCAGTATGGCACACTGAGT 5170
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 387 TCGTCTCCCCGAGATCATCCAGAGAAACCTGTGTGAGGAGAGCCTCCGACATTTGGGCCA 446
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5171 TTGTAGCACCCGAGATTGTCAATCAGAGCCCGTGTCTGAGTCACTGACATCTGGCCTG 5230
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 447 TGGGTGTCTCTCTACTCTGACCTGACCTCTCATCCCATTTGCCGCGAGAGTGACC 506
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5231 TGGGTGTGTGGCTTCTCTGTCTGACAGGAATCTCCCCCTTTGTTGGGGAATGACC 5290
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 507 GTGCCACCTCTCTGAACGTCTCTGGAGGGCGCGTGTCTATGGAGCAGCCCCCATGGCTGCC 566
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5291 GGACAAATTTGATGAACATCCGAACTACAACGTGGCTTCGAGGAGACCACATTCCTGA 5350
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 567 ACCTCAGGAAAGACGCCAAAGACTTTCATCAAGGCTACGCTGCAGAGAGCCCTCAGGCC 626
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5351 GCCTGAGCAGGGAGGCCCGGGCTTCTCTCATCAAGTGTGGTCAGGACCGGCTG--A 5407
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 627 GGCCTAGTGGGCCAGTGCTCTCCACCCCTGTTCTTGAATCCATGCTGGGAGG 686
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5408 GACCTACCGCAGAAAGACCCCTAGAATCCTTGGTT-----CAAACTCAGGCAA 5458
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 687 AGGCCACTTTCATCAACACCAAGCAGCTCAAGTTCTCTCGCCCGAAGTCGCTGGCAGC 746
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5459 AGGGCGCAGAGTGAGCACGGATCACCTGAAGCTATTCTCTCCGGCGGAGGTGGCAGC 5518
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 747 GTTCCCTGATGAGCTACAAGTCCATCTGTGTGATGCGCTCCATCCCTGAGCTGTGGGG 806
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5519 GCTCCAGATCAGCTACAAATGCCACCTGTGTGCTGCGCCCATCCCGAGCTGCTGGGG 5578
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 807 GCCCACCGA 816
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 5579 CCCCCCAGA 5588
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

Search completed: March 20, 2005, 18:55:21
Job time : 1755.97 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:15:10 ; Search time 639.09 Seconds
(without alignments)
10198.307 Million cell updates/sec

Title: US-10-077-130-6_COPY_13600_14700

Perfect score: 1101
Sequence: 1 cacagcggccacactgtgac.....agcaggaaggcccatgttc 1101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980a:*

2: Geneseqn1990a:*

3: Geneseqn2000a:*

4: Geneseqn2001a:*

5: Geneseqn2001bs:*

6: Geneseqn2002a:*

7: Geneseqn2002bs:*

8: Geneseqn2003a:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004a:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1101	100.0	24120	8	ABX11642 Human ser
2	1076	97.7	20489	12	ADQ22881 Human ser
3	886.8	80.5	14061	6	ABV99363 Human NOV
4	886.8	80.5	14109	6	ABV99362 Human NOV
5	204.2	18.5	1494	10	ADF58317 Human pol
6	71.2	6.5	2000	8	ADA71938 Rice gene
7	53.4	4.9	2534	11	ADM03261 Human cDN
8	53.4	4.9	7564	8	ABZ24581 Human cel
9	50.2	4.6	1005	6	ABN21414 Human ORF
10	49.4	4.5	2000	8	ADA71938 Rice gene
11	48.6	4.4	1813	12	ADQ22659 Human sof
12	48.6	4.4	2167	13	ACN41796 Human dia
13	48.6	4.4	2167	13	ACN41509 Human dia
14	46.6	4.2	3613	12	ADQ21633 Human sof
15	46.6	4.2	4176	12	ADQ25448 Human sof
16	45.6	4.1	2010	1	AAN90362 Glucose i
17	45	4.1	2145	5	AAS66467 DNA encod
18	45	4.1	33529	5	AAS17367 DNA seque
19	45	4.1	67251	10	ADC26995 Sorangium
20	44.8	4.1	765	11	ABD16446 Pseudomon

c	21	44.8	4.1	1329	11	ABD16328	Abd16328 Pseudomon
c	22	44.8	4.1	5121	11	ABD16585	Abd16585 Pseudomon
c	23	44	4.0	735	11	ABD15635	Abd15635 Pseudomon
c	24	44	4.0	1410	11	ABD15421	Abd15421 Pseudomon
c	25	44	4.0	27705	10	ADC26979	Adc26979 Sorangium
c	26	43.8	4.0	815	6	AAS62129	Aas62129 Porcine m
c	27	43.8	4.0	817	6	AAS62126	Aas62126 Porcine m
c	28	43.8	4.0	819	6	AAS62125	Aas62125 Porcine m
c	29	43.8	4.0	820	6	AAS62127	Aas62127 Porcine m
c	30	43.8	4.0	825	6	AAS62128	Aas62128 Porcine m
c	31	43.8	4.0	1038	13	ADS55903	AdS55903 Bacterial
c	32	43.8	4.0	9590	8	ACA45481	ACA45481 Prokaryot
c	33	43.8	4.0	177587	11	ACN44806	Acn44806 Human gen
c	34	43.6	4.0	2737	4	AAI59251	Aai59251 Human pol
c	35	43.6	4.0	3999	4	AAI61037	Aai61037 Human pol
c	36	43.6	4.0	5378	13	ADR67197	Adr67197 Human bla
c	37	43.6	4.0	5382	10	ADD14722	Add14722 Human arc
c	38	43.4	3.9	1338	12	ADP67802	Adp67802 Rice G339
c	39	43.2	3.9	1578	12	ACH87202	Ach87202 Human gen
c	40	43.2	3.9	3015	13	ADQ86796	Adq86796 Human tum
c	41	43.2	3.9	3122	4	AAH14898	Aah14898 Human cDN
c	42	43.2	3.9	3340	5	AAS45087	Aas45087 cDNA enco
c	43	43	3.9	3131	4	ABL04502	Abi04502 Drosophil
c	44	43	3.9	3131	4	ABL20422	Abi20422 Drosophil
c	45	43	3.9	5304	4	ABL20423	Abi20423 Drosophil

ALIGNMENTS

RESULT 1
ABX11642

ID ABX11642 standard; cDNA; 24120 BP.

AC ABX11642;

DT 09-MAY-2003 (first entry)

DE Human serine/threonine or protein kinase 12599, cDNA.

KW Human; ss; Gene; serine/threonine kinase; protein kinase; 12599;
KW cardiovascular disease; heart failure; myocardial infarction;
KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma;
KW blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
KW haemolytic anaemia; cellular proliferative disorder; cancer;
KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
KW psoriasis; inflammatory bowel disease; rheumatoid arthritis;
KW multiple sclerosis.

XX Homo sapiens.

OS Homo sapiens.

Key Location/Qualifiers

5'UTR 1..71

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

DR WPI; 2003-298729/29.
XX P-PSDB; ABG76187.
PT Novel isolated human protein kinase, designated 59079 or 12599
PT polypeptide, useful as diagnostic and therapeutic agents for preventing
PT cardiovascular diseases, proliferative disorders, and protein kinase
PT disorders.
XX
PS Claim 2; Page 58-84; 119pp; English.
XX
CC The invention relates to an isolated human serine/threonine or protein
CC kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule
CC comprising at least 85% identity to the nucleic acids appearing as
CC ABX11641 and ABX11642 or their complement, a naturally occurring variant
CC of the kinases or their fragments. Also included are a non-human host
CC cell containing the nucleic acids, an antibody specific for the proteins,
CC identifying a compound which binds to the kinase (by contacting the
CC kinase or a cell expressing the kinase with a test compound and
CC determining whether the kinase binds to the test compound) and modulating
CC the activity of kinase using the identified compound. The kinases and
CC their encoding nucleic acids are useful as diagnostic and therapeutic
CC agents for preventing a disease or condition associated with an aberrant
CC or unwanted 59079 or 12599 activity in a subject, including
CC cardiovascular diseases such as heart failure, and myocardial infarction;
CC disorders involving blood vessels such as atherosclerosis, and Kaposi's
CC sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia,
CC Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders
CC such as cancer; and protein kinase disorders such as autoimmune
CC disorders, diabetes mellitus, psoriasis, inflammatory bowel disease,
CC rheumatoid arthritis, and multiple sclerosis (many examples of diseases
CC and disorders are included in the specification). The kinases, their
CC encoding nucleic acids and antibodies are useful in screening assays,
CC detection assays (e.g. forensic biology), and predictive medicine (e.g.
CC diagnostic assays, prognostic assays, and monitoring clinical trials and
CC pharmacogenomics). The kinases and their encoding nucleic acids are
CC useful as query sequences to perform a search against public databases to
CC identify other family members or related sequences. The present sequence
CC encodes the kinase 12599
XX
SQ Sequence 24120 BP; 4719 A; 7265 C; 8278 G; 3858 T; 0 U; 0 Other;

Query Match 100.0%; Score 1101; DB 8; Length 24120;
Best Local Similarity 100.0%; Pred. No. 5.5e-233;
Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CACAGACCCACACTGTGACACTGTCTTGGGACGCTCCCATGATGATGGAGCGGTGGT 13730
DB CACAGACCCACACTGTGACACTGTCTTGGGACGCTCCCATGATGATGGAGCGGTGGT 13731

QY 61 CTCTGTGGCTACCGGTGGAGGTGAGAGAGGGGGCCACAGGCCAGTGGCGGCTGTGCCAC 120
DB CACAGACCCACACTGTGACACTGTCTTGGGACGCTCCCATGATGATGGAGCGGTGGT 13731

QY 121 GAGCTGTGTGCTGACCCGAGTGTGTGTGATGCTGCTGGCCAGTGGCGGCTGTGCCAC 180
DB GAGCTGTGTGCTGACCCGAGTGTGTGTGATGCTGCTGGCCAGTGGCGGCTGTGCCAC 13791

QY 181 TTCGCTGTGGCAGCTGTGGGCCCTGTGGTGTCTGGGAAACCGGTTACCTGCCCCAGACA 240
DB TTCGCTGTGGCAGCTGTGGGCCCTGTGGTGTCTGGGAAACCGGTTACCTGCCCCAGACA 13851

QY 241 GTGGGGCTTGCAGAGCCACCGAGCCTGTGCTCTCCAGCCCTCAGCCCTGTAGAGCCGG 300
DB GTGGGGCTTGCAGAGCCACCGAGCCTGTGCTCTCCAGCCCTCAGCCCTGTAGAGCCGG 13911

QY 301 CAGTGTGGCAGCTGTGAGATGTCTCTCTGAGGCTTGAGGTGGTGTGGTGTGGTGTAG 360
DB CAGTGTGGCAGCTGTGAGATGTCTCTCTGAGGCTTGAGGTGGTGTGGTGTGGTGTAG 13971

QY 361 GTCATCTGGCCACAGGGAATGAGCGCATCCAGCCCGTGGCGGTTTCGAGGTGGTCTCC 420
DB GTCATCTGGCCACAGGGAATGAGCGCATCCAGCCCGTGGCGGTTTCGAGGTGGTCTCC 14031

QY 421 CAGGGTCGGCAACAGATGCTGTGATCAAGGGCTTTCACGGCAGAAAGACCCAGGCGAGTAC 480
DB CAGGGTCGGCAACAGATGCTGTGATCAAGGGCTTTCACGGCAGAAAGACCCAGGCGAGTAC 14091

QY 481 CACTGTGTGGCTGTGCTCAGGGCTCCATCTGCGCTTCCGCTGCTGCTTCCAGGTGGCACTG 540
DB CACTGTGTGGCTGTGCTCAGGGCTCCATCTGCGCTTCCGCTGCTGCTTCCAGGTGGCACTG 14151

QY 541 AGCCCAAGCCTCTGTGGATGAGGCCCTTCAGCCCAAGCTTGGCCCGGAGGAGCCAGGAG 600
DB AGCCCAAGCCTCTGTGGATGAGGCCCTTCAGCCCAAGCTTGGCCCGGAGGAGCCAGGAG 14211

QY 601 GGTGACCTGCACTACTGTGTGGAGGCCCTGTGCTCGGAAACGTTCGATGAGCCGTGAGCCC 660
DB GGTGACCTGCACTACTGTGTGGAGGCCCTGTGCTCGGAAACGTTCGATGAGCCGTGAGCCC 14271

QY 661 ACGTGTGACTCCATTAGCAGCTGTCAGAGAGGAGCGGCCCTCGCAGCGCTTCCACAG 720
DB ACGTGTGACTCCATTAGCAGCTGTCAGAGAGGAGCGGCCCTCGCAGCGCTTCCACAG 14331

QY 721 GAGGCAGAGAGGTGGCACCTGATCTCTCTGAGGGCTACTCCAGCGCGGATGAGCTGGCC 780
DB GAGGCAGAGAGGTGGCACCTGATCTCTCTGAGGGCTACTCCAGCGCGGATGAGCTGGCC 14391

QY 781 CGCATCTGGAGATGCTGACCTCTTCACACACAGCTCTCTGATGATGATGATGATGATGAT 840
DB CGCATCTGGAGATGCTGACCTCTTCACACACAGCTCTCTGATGATGATGATGATGATGAT 14451

QY 841 CTTTCCCTGTGCTACCTTACCTTCAAGAGGCTGGAGGCTGGAGGCTGCTTCCACCTGCGCCAG 900
DB CTTTCCCTGTGCTACCTTACCTTCAAGAGGCTGGAGGCTGGAGGCTGCTTCCACCTGCGCCAG 14511

QY 901 AAGGTTGGGGCCCCCAGCAGCCCCCTCTGTGAAGCCACAGCAGCAGCAGCAGCAGCAGCTGGCT 960
DB AAGGTTGGGGCCCCCAGCAGCCCCCTCTGTGAAGCCACAGCAGCAGCAGCAGCAGCAGCTGGCT 14571

QY 961 GCTGTGCGCCACACCTGCGGAGACCTGAGCACCACAAAGACCTGGGTGATCTTCAATGGAC 1020
DB GCTGTGCGCCACACCTGCGGAGACCTGAGCACCACAAAGACCTGGGTGATCTTCAATGGAC 14631

QY 1021 AAGGCAGCTGTGAAGATCCAGGCTGCTTTAAGGGCTTCAAGGCTCCGGAAGGATGAAG 1080
DB AAGGCAGCTGTGAAGATCCAGGCTGCTTTAAGGGCTTCAAGGCTCCGGAAGGATGAAG 14691

QY 1081 CAGCAGAGAGGGGCCCATGTTTC 1101
DB CAGCAGAGAGGGGCCCATGTTTC 14751

RESULT 2
ADQ22881

ID ADQ22881 standard; DNA; 20489 BP.

XX AC ADQ22881;

XX DT 26-AUG-2004 (first entry)

XX DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5701.

XX KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW db.

XX OS Homo sapiens.

XX PN WO2004048938-A2.

XX PD 10-JUN-2004.

XX PF 26-NOV-2003; 2003WO-US038193.

XX PR 26-NOV-2002; 2002US-0429739P.

XX PA (PROT-) PROTEIN DESIGN LABS INC.

XX PI Aziz N, Gineburg WM, Zlotnik A;
XX DR WPI; 2004-441208/41.
XX PT Early detection of soft tissue sarcoma comprises determining expression
XX PT of a gene in a first soft tissue sample and a normal soft tissue sample
XX PT and comparing the gene expression, also useful in treating soft tissue
XX PT sarcoma.
XX PS Example 2; SEQ ID NO 5701; 210pp; English.
XX CC The invention relates to a novel method for detecting soft tissue sarcoma
XX CC which comprises obtaining a first soft tissue sample from an individual
XX CC and a normal soft tissue sample from the same or different individual,
XX CC determining the expression of a gene in both samples and comparing the
XX CC expression of the gene in both soft tissue samples, where a higher level
XX CC of protein expression in the first soft tissue sample indicates the
XX CC presence of soft tissue sarcoma. The method of the invention has
XX CC cytostatic applications and may be useful for detecting soft tissue
XX CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX CC acid sequences may be useful in diagnostic and screening applications.
XX CC The current sequence is that of a human soft tissue sarcoma-upregulated
XX CC DNA of the invention. The current sequence is not shown within the
XX CC specification per se but was submitted in CD format by the inventor.
XX SQ Sequence 20489 BP; 3917 A; 5815 C; 6983 G; 3225 T; 0 U; 549 Other;
Query Match 97.7%; Score 1076; DB 12; Length 20489;
Best Local Similarity 97.7%; Pred. No. 1.8e-227;
Matches 1076; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
QY 1 CACAGCAGCCACTGTGACACTGTCTTGGGAGCTCCCATGATGATGGAGCGGTGT 60
DB 13677 CACAGCAGCCACTGTGACACTGTCTTGGGAGCTCCCATGATGATGGAGCGGTGT 13736
QY 61 CTCTGTGCTACCGGTGGAGGTGAAGAGGGGGCCACAGGCCAGTGGCGGTGGCCAC 120
DB 13737 CTCTGTGCTACCGGTGGAGGTGAAGAGGGGGCCACAGGCCAGTGGCGGTGGCCAC 13796
QY 121 GAGCTGGTGGTGGACCCGAGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
DB 13797 GAGCTGGTGGTGGACCCGAGTGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 13856
QY 181 TTCCGTGTGGCAGCTGTGGGCGCTGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 240
DB 13857 TTCCGTGTGGCAGCTGTGGGCGCTGTGGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 13916
QY 241 GTGGCGCTTGACAGCCACCGAGCTGTGCTTCCCGAGCCCTGAGCCCTGAGAGCCGG 300
DB 13917 GNN 13976
QY 301 CAGGTGGCAGCTGGTGAAGATGTCTCTGGAGCTTGAAGTGGTGGTGGTGGTGGTGGTGGTGG 360
DB 13977 CAGGTGGCAGCTGGTGAAGATGTCTCTGGAGCTTGAAGTGGTGGTGGTGGTGGTGGTGG 14036
QY 361 GTCAATCTGGCAAGGGAATGAGCGCATCCAGCCCGGTGGGCGGTTCAGAGTGGTCTCC 420
DB 14037 GTCAATCTGGCAAGGGAATGAGCGCATCCAGCCCGGTGGGCGGTTCAGAGTGGTCTCC 14096
QY 421 CAGGTGGCAGCATGCTGGTGTATCAAGGCTTCAGCGCAGAGACCGAGGCGAGTAC 480
DB 14097 CAGGTGGCAGCATGCTGGTGTATCAAGGCTTCAGCGCAGAGACCGAGGCGAGTAC 14156
QY 481 CACTGTGGCTGGCTCAGGGCTCCATCTGCGCTGGGCTGGCCACCTTCCAGGTGGCACTG 540
DB 14157 CACTGTGGCTGGCTCAGGGCTCCATCTGCGCTGGGCTGGCCACCTTCCAGGTGGCACTG 14216
QY 541 AGCCAGCGCTCTGTGGATGAGGCCCTCAGCCAGCTTTCGCCCGCCAGGCGAGCCAGGAG 600
DB 14217 AGCCAGCGCTCTGTGGATGAGGCCCTCAGCCAGCTTTCGCCCGCCAGGCGAGCCAGGAG 14276
QY 601 GGTGACCTGCACCTACTGTGGAGGCGCTTGGCTCGGAAACGTGCGCATGAGCGGTGAGGCC 660

DB 14277 GGTGACCTGCACCTACTGTGGAGGCTTGGCTCGGAAACGTGCGCATGAGCGGTGAGGCC 14336
QY 661 ACCTTGGACTCCATTAGCGAGCTGCAGAGGAGGAGCGCGCTCGCAGCGCTGCACAG 720
DB 14337 ACCTTGGACTCCATTAGCGAGCTGCAGAGGAGGAGCGCGCTCGCAGCGCTGCACAG 14396
QY 721 GAGCAGAGGAGGTGGCAGCTGTCTCTGAAGGCTACTCCAGCGCGCATGAGTGGCC 780
DB 14397 GAGCAGAGGAGGTGGCAGCTGTCTCTGAAGGCTACTCCAGCGCGCATGAGTGGCC 14456
QY 781 CGCACTGGAGATGCTGACCTCTCACACACAGCTCTGATGATGAGTCCCGGCGAGCCACC 840
DB 14457 CGCACTGGAGATGCTGACCTCTCACACACAGCTCTGATGATGAGTCCCGGCGAGCCACC 14516
QY 841 CTTTCCCTGGTCACTACCTCAAGAAGGCTGGGAGGCCAGGACCTCACCACTGGCCAGC 900
DB 14517 CTTTCCCTGGTCACTACCTCAAGAAGGCTGGGAGGCCAGGACCTCACCACTGGCCAGC 14576
QY 901 AAGTTGGGCGCCAGCAGCCCTCTGTGAAGCCACAGCAGCAGCAGGAGCCACTGGCT 960
DB 14577 AAGTTGGGCGCCAGCAGCCCTCTGTGAAGCCACAGCAGCAGCAGGAGCCACTGGCT 14636
QY 961 GCTGTGCCCGCCACCACTGGGAGACTGAGCACCAGAACCTGGGTGATCCCTCAATGAC 1020
DB 14637 GCTGTGCCCGCCACCACTGGGAGACTGAGCACCAGAACCTGGGTGATCCCTCAATGAC 14696
QY 1021 AAGCAGCTGTGAAGATCCAGGCTGCCTTTAAGGCTCAAGGTCGCGAAGAGATGAAG 1080
DB 14697 AAGCAGCTGTGAAGATCCAGGCTGCCTTTAAGGCTCAAGGTCGCGAAGAGATGAAG 14756
QY 1081 CAGCAGGAAGGCGCCATGTT 1101
DB 14757 CAGCAGGAAGGCGCCATGTT 14777
RESULT 3
ID ABV99363 standard; DNA; 14061 BP.
XX AC ABV99363;
XX DT 27-JAN-2003 (first entry)
XX DE Human NOV13b coding sequence.
XX KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
XX KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
XX KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
XX KW antiinfertility; cerebroprotective; gene therapy; NOVA; NOV; fertility;
XX KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
XX KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
XX KW immune disorder; haematopoietic disorder; cardiovascular disorder;
XX KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
XX KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
XX OS Homo sapiens.
XX PN WO200272771-A2.
XX PD 19-SEP-2002.
XX PF 08-MAR-2002; 2002WO-US007288.
XX PR 08-MAR-2001; 2001US-0274101P.
XX PR 08-MAR-2001; 2001US-0274194P.
XX PR 08-MAR-2001; 2001US-0274281P.
XX PR 09-MAR-2001; 2001US-0274322P.
XX PR 12-MAR-2001; 2001US-0274849P.
XX PR 13-MAR-2001; 2001US-0275235P.
XX PR 13-MAR-2001; 2001US-0275579P.

PR 13-MAR-2001; 2001US-0275601P.
 PR 14-MAR-2001; 2001US-0276000P.
 PR 16-MAR-2001; 2001US-0276776P.
 PR 19-MAR-2001; 2001US-0276994P.
 PR 20-MAR-2001; 2001US-0277329P.
 PR 20-MAR-2001; 2001US-0277321P.
 PR 20-MAR-2001; 2001US-0277327P.
 PR 20-MAR-2001; 2001US-0277338P.
 PR 21-MAR-2001; 2001US-0277791P.
 PR 22-MAR-2001; 2001US-0277833P.
 PR 23-MAR-2001; 2001US-0278152P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0278999P.
 PR 27-MAR-2001; 2001US-0279036P.
 PR 28-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0279995P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 02-APR-2001; 2001US-0280900P.
 PR 04-APR-2001; 2001US-0281194P.
 PR 13-APR-2001; 2001US-0283675P.
 PR 30-APR-2001; 2001US-0287424P.
 PR 02-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 03-MAY-2001; 2001US-0288528P.
 PR 15-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 16-MAY-2001; 2001US-0291240P.
 PR 30-MAY-2001; 2001US-0294485P.
 PR 31-MAY-2001; 2001US-0294889P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 18-OCT-2001; 2001US-0325681P.
 PR 31-OCT-2001; 2001US-0330380P.
 PR 14-NOV-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332727P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 21-NOV-2001; 2001US-0333272P.
 PR 03-DEC-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337426P.
 PR 04-DEC-2001; 2001US-0338092P.
 PR 04-DEC-2001; 2001US-0337185P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 08-MAR-2002; 2002US-00093463.
 PR (CURA-) CURAGEN CORP.
 PR Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 PI Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
 PI Pena CE, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 PI Taupier RJ, Fadigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 PI Zhong M;
 XX WPI; 2002-732824/79.
 DR P-PSDB; ABP70085.
 XX
 PT New NOVX polypeptides and polynucleotides, useful for preventing,
 PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
 PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 PT disorders, and asthma.
 XX

PS Claim 16; Page 138-142; 619pp; English.
 XX The present invention relates to new isolated proteins (NOVX) and their
 CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
 CC any number from 1 to 48. The NOVX proteins and coding sequences are
 CC useful in the manufacture of a medicament for treating a syndrome
 CC associated with a human disease, preferably a NOVX-associated disorder.
 CC The NOVX coding sequences and proteins are useful for treating,
 CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
 CC obesity, infectious disease, anorexia, cancer-associated cachexia,
 CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 CC disease, immune disorders, hematopoietic disorders, cardiovascular
 CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 CC disturbances associated with obesity, metabolic syndrome X or wasting
 CC disorders associated with chronic diseases or various cancers. The NOVX
 CC coding sequences and proteins may also be used as targets for the
 CC identification of small molecules that modulate or inhibit e.g.
 CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 CC wound healing and angiogenesis, in gene therapy, in generation of
 CC antibodies that bind immunospecifically to NOVX substances for use in
 CC therapeutic or diagnostic methods
 XX
 SQ Sequence 14061 BP; 2854 A; 4029 C; 4839 G; 2339 T; 0 U; 0 Other;
 Query Match 80.5%; Score 886.8; DB 6; Length 14061;
 Best Local Similarity 98.9%; Pred. No. 9.2e-186;
 Matches 904; Conservative 0; Mismatches 7; Indels 3; Gaps 1;
 QY 1 CACAGAGCCACACTGTGACACTGTCTTGGGAGCTCCCATGAGTGTGAGGCGGTGGT 60
 DB 9366 CGCAGAGCCACACTGTGACACTGTCTTGGGAGCTCCCATGAGTGTGAGGCGGTGGT 9425
 QY 61 CTCTGTGGCTACCGCGTGGAGGTGAAGAGCGGCGCCACAGCCAGTGGCGGCTGTGCCAC 120
 DB 9426 CTCTGTGGCTACCGCGTGGAGGTGAAGAGCGGCGCCACAGCCAGTGGCGGCTGTGCCAC 9485
 QY 121 GAGCTGTGCTGGACCCGAGTGTGTGGTGGTGGCTGGCCCGCGGGAGAGACTACCGC 180
 DB 9486 GAGCTGTGCTGGACCCGAGTGTGTGGTGGTGGCTGGCCCGCGGGAGAGACTACCGC 9545
 QY 181 TTCGGTGTGGCAGCTGTGGGCGCTGTGGGTGTGGGAGACCGGTTTACCTGCCCGACACA 240
 DB 9546 TTCGGTGTGGCAGCTGTGGGCGCTGTGGGTGTGGGAGACCGGTTTACCTGCCCGACACA 9605
 QY 241 GTGGCGCTTGCAGAGCCACCGAAGCTGTGCTCTCCAGCCCTCAGCCCTCAGAGCGCG 300
 DB 9606 GTGGCGCTT---GAGCCACCGAAGCTGTGCTCTCCAGCCCTCAGCCCTCAGAGCGCG 9662
 QY 301 CAGGTGGCAGCTGGTGAAGATGTCTCTGGAGCTTGAAGTGTGGTGTGGTGTGGTGTGG 360
 DB 9663 CAGGTGGCAGCTGGTGAAGATGTCTCTGGAGCTTGAAGTGTGGTGTGGTGTGGTGTGG 9722
 QY 361 GTCATCTGGCACAAAGGGAATGGAGCGCATCCAGCCCGTGGGCGGTTCCAGGTGTCTCC 420
 DB 9723 GTCATCTGGCACAAAGGGAATGGAGCGCATCCAGCCCGTGGGCGGTTCCAGGTGTCTCC 9782
 QY 421 CAGGTCGGCACAAGATGTGTGTGATCAAGGGGTTTCAAGGCAAGAACACAGGCGGAGTAC 480
 DB 9783 CAGGTCGGCACAAGATGTGTGTGATCAAGGGGTTTCAAGGCAAGAACACAGGCGGAGTAC 9842
 QY 481 CACTGTGGCTGGCTCAGGGCTCCATCTGCTGCTGGGCTGGCCACCTTCAGAGTGGCACTG 540
 DB 9843 CACTGTGGCTGGCTCAGGGCTCCATCTGCTGCTGGGCTGGCCACCTTCAGAGTGGCACTG 9902
 QY 541 AGCCAGGCTCTGTGGATGAGGCGCTCAGCCCGTGGCTGGCCCGGAGCAGCCAGGAG 600
 DB 9903 AGCCAGGCTCTGTGGATGAGGCGCTCAGCCCGTGGCTGGCCCGGAGCAGCCAGGAG 9962
 QY 601 GGTGACCTGCACCTACTGTGGAGGCGCTGGGTTCGGAACGTCGCAATGAGCCGTGAGCCC 660
 DB 9963 GGTGACCTGCACCTACTGTGGAGGCGCTGGGTTCGGAACGTCGCAATGAGCCGTGAGCCC 10022
 QY 661 ACCTGGACTCCATTAGCGAGCTGCCAGAGGAGGCGGCTCGCAGCGCTGCCACAG 720

CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods
XX
SQ Sequence 14109 BP; 2862 A; 4045 C; 4854 G; 2348 T; 0 U; 0 Other;
Query Match 80.5%; Score 886.8; DB 6; Length 14109;
Best Local Similarity 98.9%; Pred. No. 9.2e-186;
Matches 904; Conservative 0; Mismatches 7; Indels 3; Gaps 1;
QY 1 CACAGCAGCCACACTGTGACACTGTCTTGGGAGCTCCCATGATGATGAGGCGGTGGT 60
Db |||||
QY 9366 CGCAGCAGCCACACTGTGACACTGTCTTGGGAGCTCCCATGATGATGAGGCGGTGGT 9425
Db |||||
QY 61 CTCCTGTGGCTACCGGTGAGGTGAAGAGAGGGGGCCACAGGCCAGTGGCGCTGTGCCAC 120
Db |||||
QY 9426 CTCCTGTGGCTACCGGTGAGGTGAAGAGAGGGGGCCACAGGCCAGTGGCGCTGTGCCAC 9485
Db |||||
QY 121 GAGCTGTGTGCTGGACCCGAGTGTGTGTGATGCTGGCCCGGGGAGACCTACCGC 180
Db |||||
QY 9486 GAGCTGTGTGCTGGACCCGAGTGTGTGTGATGCTGGCCCGGGGAGACCTACCGC 9545
Db |||||
QY 181 TTCCTGTGGCAGCTGTGGGCTCTGTGGTGTCTGGGAAACCGTTTCACTTCCGCCAGACA 240
Db |||||
QY 9546 TTCCTGTGGCAGCTGTGGGCTCTGTGGTGTCTGGGAAACCGTTTCACTTCCGCCAGACA 9605
Db |||||
QY 241 GTGGGCTTGCAGAGCCACGAGCTGTGCTTCCCGAGCCCTCAGCCCTGAGAGCCGG 300
Db |||||
QY 9606 GTGGGCTT---GAGCCACCGAAGCTGTGCTTCCCGAGCCCTCAGCCCTGAGAGCCGG 9662
Db |||||
QY 301 CAGGTGGCAGCTGTGTAAGATGTCTCTGTGAGCTTTGAGGTGGTGGCTGAGGCTGTGAG 360
Db |||||
QY 9663 CAGGTGGCAGCTGTGTAAGATGTCTCTGTGAGCTTTGAGGTGGTGGCTGAGGCTGTGAG 9722
Db |||||
QY 361 GTCACTTGGCACAAGGGAATGAGCGCATCAGCCCGTGGCGGTTGAGGTGTCTCC 420
Db |||||
QY 9723 GTCACTTGGCACAAGGGAATGAGCGCATCAGCCCGTGGCGGTTGAGGTGTCTCC 9782
Db |||||
QY 421 CAGGTGGCAGCTGTGTAAGATGTCTCTGTGAGCTTTGAGGTGGTGGCTGAGGCTGTGAG 480
Db |||||
QY 9783 CAGGTGGCAGCTGTGTAAGATGTCTCTGTGAGCTTTGAGGTGGTGGCTGAGGCTGTGAG 9842
Db |||||
QY 481 CACTGTGGCTTGGCTCAGGCTTCCATCTGCGCTGCGCTGCGCTTCCAGGTGGCACTG 540
Db |||||
QY 9843 CACTGTGGCTTGGCTCAGGCTTCCATCTGCGCTGCGCTGCGCTTCCAGGTGGCACTG 9902
Db |||||
QY 541 AGCCGAGCTCTGTGGATGAGGCTTCCAGCCAGCTTGGCCCGGAGGCAAGCCAGGAG 600
Db |||||
QY 9903 AGCCGAGCTCTGTGGATGAGGCTTCCAGCCAGCTTGGCCCGGAGGCAAGCCAGGAG 9962
Db |||||
QY 601 GGTGACCTGCACTTCTGTGGAGGCTTGGCTGGAACCTTCCATGAGCCGTGAGCC 660
Db |||||
QY 9963 GGTGACCTGCACTTCTGTGGAGGCTTGGCTGGAACCTTCCATGAGCCGTGAGCC 10022
Db |||||
QY 661 ACCTGTGACTCCATTAGCAGCTGCGAGGAGGAGCGGCTGCGAGCGCTTCCACAG 720
Db |||||
QY 10023 ACCTGTGACTCCATTAGCAGCTGCGAGGAGGAGCGGCTGCGAGCGCTTCCACAG 10082
Db |||||
QY 721 GAGCAGAGGAGGTGGACCTGATCTCTTCTGAAAGGCTACTCCAGCGCGATGAGCTGGCC 780
Db |||||
QY 10083 GAGCAGAGGAGGTGGACCTGATCTCTTCTGAAAGGCTACTCCAGCGCGATGAGCTGGCC 10142
Db |||||
QY 781 CGCAGTGGAGATGCTGACCTCTCACACAGCTCTGATGATGATGATGATGATGATGATGATGAT 840
Db |||||
QY 10143 CGCAGTGGAGATGCTGACCTCTCACACAGCTCTGATGATGATGATGATGATGATGATGATGAT 10202
Db |||||
QY 841 CCTTCCCTGTGTCACTTCAAGAGGCTGGGAGGCGAGGACCTTCAACCACTTGGCCAGC 900
Db |||||
QY 10203 CCTTCCCTGTGTCACTTCAAGAGGCTGGGAGGCGAGGACCTTCAACCACTTGGCCAGC 10262
Db |||||
QY 901 AAGTTTGGGGCCCC 914

Db 10263 AAGTTGAGCCCCC 10276
RESULT 5
ADF58317
ID ADF58317 standard; cDNA; 1494 BP.
XX
AC ADF58317;
XX
DT 12-FEB-2004 (first entry)
XX
DE Human polynucleotide sequence SEQ ID NO:684.
XX
KW biological activity; genetic engineering; hybridisation probe; oligomer;
KW primer; chromosome mapping; gene mapping; recombinant protein production;
XX human; gene; ss.
XX OS Homo sapiens.
XX WO2003080795-A2.
XX PD 02-OCT-2003.
XX PF 09-AUG-2002; 2002WO-US025485.
XX PR 09-AUG-2001; 2001US-0311261P.
XX (HYSE-) HYSEQ INC.
XX PA Tang YT, Yang Y, Wang Z, Weng G, Ma Y;
XX PT WPI: 2003-876918/81.
XX P-PSDB; ADF59317.
XX
PT New polynucleotides, useful as hybridization probes, oligomers or
PT primers, for chromosome or gene mapping, for the recombinant production
PT of proteins, and for generating antisense DNA or RNA.
XX
PS Claim 1; SEQ ID NO 684; 571pp; English.
XX
CC The present sequence represents an isolated polynucleotide sequence (I)
CC from the present invention, which encodes a polypeptide (II) with
CC biological activity. Also described: (1) a vector comprising (I); (2) an
CC expression vector comprising (I); (3) a host cell genetically engineered
CC to comprise (I) which is operatively associated with a regulatory
CC sequence that modulates expression of (I) in the host cell; (4) a
CC polypeptide (II) encoded by (I); (5) a composition comprising the
CC polypeptide of (4) and a carrier; (6) an antibody directed against the
CC polypeptide of (4); (7) detecting (I) or the polypeptide of (4) in a
CC sample; (8) identifying a compound that binds to the polypeptide of (4);
CC (9) producing the polypeptide of (4); and (10) a collection of
CC polynucleotides comprising at least one of the polynucleotide sequences
CC (I). The polynucleotides (I) can be used as hybridisation probes,
CC oligomers or primers, for chromosome or gene mapping, for the recombinant
CC production of proteins, and for generating antisense DNA or RNA.
XX
SQ Sequence 1494 BP; 317 A; 421 C; 441 G; 315 T; 0 U; 0 Other;
Query Match 18.5%; Score 204.2; DB 10; Length 1494;
Best Local Similarity 96.3%; Pred. No. 2.5e-35;
Matches 209; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 885 CTCACCACTGGCCAGCAAGTTGGGGCCCGCCAGCAGCCCTCTGTGAAGCCACAGCA 944
Db 12 CAACAACCTGTTTCTCTGTTGGGGCCCGCCAGCAGCCCTCTGTGAAGCCACAGCA 71
QY 945 GCAGGAGCCACTGGCTGCTGTGCGCCACACCTGGGAGACCTGAGCACCAGACCTGGG 1004
Db |||||
QY 72 GCAGGAGCCACTGGCTGCTGTGCGCCACACCTGGGAGACCTGAGCACCAGACCTGGG 131
Db |||||
QY 1005 TGATCCCTCAATGGAACAAGCGAGCTGTGAAGATCCAGGCTGCTTTAAGGCTTACAGGT 1064
Db |||||

Db 132 TGATCCCTCAATGGACAGGAGCTGTGAGATCCAGGCTGCTTTAAGGGCTACAAGGT 191

Qy 1065 CCGGAGGAGATGAAGCAGCAGGAGGAGGCGCCATGTTTC 1101

Db 192 CCGGAGGAGATGAAGCAGCAGGAGGAGGCGCCATGTTTC 228

RESULT 6

ADA71938

ID ADA71938 standard; DNA; 2000 BP.

XX

XX ADA71938;

XX

XX 20-NOV-2003 (first entry)

XX

XX Rice gene, SEQ ID 5263.

DE

XX Plant; bacterial infection; fungal infection; viral infection; rice;

KW gene; ds.

XX

XX Oryza sativa.

XX

XX WO200300898-A1.

XX

XX 03-JAN-2003.

XX

XX 22-JUN-2001; 2001WO-IB001105.

XX

XX 22-JUN-2001; 2001WO-IB001105.

XX

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX

XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX

XX WPI; 2003-175290/17.

XX

PT Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to

PT bacterial, fungal or viral infection by determining or detecting plant

PT gene expression.

XX

XX Claim 27; SEQ ID NO 5263; 899pp; English.

XX

XX The present invention relates to a method (M1) for identifying genes

CC involved in plant resistance or response to pathogenic infection. M1

CC comprises identifying a gene whose expression is significantly altered in

CC the incompatible interaction of plant gene expression relative to

CC expression of the gene in an uninfected plant, in a mutant plant that

CC does not express a gene associated with response to pathogenic infection,

CC or in a corresponding incompatible or compatible interaction. (M1) is

CC useful for conferring resistance to resistance or tolerance to a plant to

CC bacterial, fungal or viral infection. The present sequence was used to

XX illustrate the invention.

XX

SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 6.5%; Score 71.2; DB 8; Length 2000;

Best Local Similarity 9.4%; Pred. No. 6e-06;

Matches 66; Conservative 340; Mismatches 288; Indels 6; Gaps 2;

Qy 328 CTGAGCTTGAGTGGCTGAGCTGGTGTGATCTGTCACAGGAGGAGTGGAGCGC 387

Db 78 SKWSSGSGRGMKRYKRSRWRGRGRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMR 137

Qy 388 ATCCAGCCGGTGGGCGGTGAGTGGTCTCCAGGCTGCGCAACAGATGCTGTGATC 447

Db 138 GKSRSYMWYARCGGCKRKSXGSGWGTCTRRGAGGGSWSGATKSGSMRMM 197

Qy 448 AAGGCTTTCAGCGCAGAGACAGCGGCGAGTACCACTGTGGCTGGCTCAGGCTCCATC 507

Db 198 SSCGRSGRRSAYSRYYGTSRKYTKYKMTYYSASRCMRVMTTYSWACSSYTWCRSK 257

Qy 508 TGCCCTGGCTGCCACCTTCCAGGTGGCACTGAGCCAGCCCTCT--GTGATGAGGCC 564

Db 258 RRSMMWKMWRKWSRSYSGWYSWSYKSYYSRMCYMYRGGWGWRATRYWGRG 317

Qy 565 CCTCAGCCAGCTTGCCTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 624

Db 318 YMSRWMMYKMYVRYGKMGKRWAGMMRSMCRSKACYYMRWRMRTRRRWAK 377

Qy 625 CCCTTGCTCGAAACG---TCGCAATGAGCGGTGAGCCACGCTGGACTCCATGAGGAG 681

Db 378 KSSRTSRKKEKWCWKRYKMRGYSRMRSCRRARMMKRCRSGRAWKMGCRGCMTCRMK 437

Qy 682 CTGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 741

Db 438 SYGMMRWKSKRSMASRYKMRSMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMR 497

Qy 742 GATCTCTGAAGGCTACTCCACGCGCCGAGCTGAGCTGGCCGCGCAGCTGGAGAGATGCTGACCTC 801

Db 498 RKRRRRRGRMR 557

Qy 802 TCACACACAGCTCTGATGATGAGTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 861

Db 558 RYKRMYYKMMWYKRYKSGSWYKMSYASCMKSAKAGAKMKRSMKSAWSKSMRSS 617

Qy 862 AAGAAGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 921

Db 618 RKRCRCKASRSRSMKRYAMMGMTSGSRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMR 677

Qy 922 CCCTCTGTGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 981

Db 678 KYGSYRYRYRAWCYMR 737

Qy 982 GACTTGAGCAGCAAGACCTGGGTGATCCCTCAATGGACA 1021

Db 738 KCSWKYSMMYYWSMMWAKTWKMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMRMR 777

RESULT 7

ADM03261

ID ADM03261 standard; cDNA; 2534 BP.

XX

XX ADM03261;

XX

XX 20-MAY-2004 (first entry)

XX

XX Human cDNA of the invention SEQ ID NO:1946.

DE

XX ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.

KW

XX Homo sapiens.

OS

XX EP1347046-A1.

XX

XX 24-SEP-2003.

PD

XX

XX 12-APR-2002; 2002EP-00008400.

PF

XX

XX 22-MAR-2002; 2002JP-00137785.

PR

XX

XX (REAS-) RES ASSOC BIOTECHNOLOGY.

PA

XX

XX Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;

PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;

PI Seki N, Yoshikawa T, Otsuka M, Negahari K, Masuho Y;

XX

XX WPI; 2003-723558/69.

DR

XX P-PsDB; ADM05704.

XX

PT New polynucleotides and polypeptides are useful in gene therapy, for

PT developing a diagnostic marker or medicines for regulating their

PT expression and activity, or as a target of gene therapy.

XX

PS Claim 1; SEQ ID NO 1946; 305pp; English.

The invention relates to a novel human polynucleotide and the encoded polypeptide. A polynucleotide of the invention may have a use in gene therapy. An oligonucleotide of the invention ADOM6202-ADOM6773 is useful as a primer for synthesizing the polynucleotide or as a probe for detecting the polynucleotide. The polynucleotides ADOM1316-ADOM3758 are useful in gene therapy, for developing a diagnostic marker or medicines for regulating their expression and activity, or as a target of gene therapy. The proteins ADOM3759-ADOM6201 encoded by the polynucleotides are useful as pharmaceutical agents. The present sequence represents a cDNA sequence of the invention.

Sequence 2534 BP; 551 A; 679 C; 778 G; 526 T; 0 U; 0 Other;

Query Match	4.9%;	Score 53.4;	DB 11;	Length 2534;
Best Local Similarity	53.6%;	Pred. No. 0.053;		
Matches 111;	Conservative 0;	Mismatches 96;	Indels 0;	Gaps 0;
Qy	287	CCCCGTGAGAGCGCGGAGTGCAGCTGGTGAAGATGCTCTCTCGAGCTTGAGGTGGTGG	346	
Db	803	CCCTGGGAAGACCATGGGTGGCCGACGGGGAGGACGTGGAGCTGGCTGTGAGCTGTAC	862	
Qy	347	CTGAGGCTGGTGAAGTCAATCTGGCACAAAGGGAATGGAGCGCATCCAGCCGGTGGCGGT	406	
Db	863	GGGCGGGAACGCCCGTGCATCTGGCTGAAGACAGGAAGGCCATCCGCAAGACCAAGT	922	
Qy	407	TCGAGGTGGTCTCCACGGTCGGCAACAGATGCTGGTGATCAAGGGCTTTCACGCGCAGAG	466	
Db	923	ATGATGTGTCTGCGAGGGCAGATGGCCATGCTGGTCAATCCGCGGGGCTCTCGTCAAGG	982	
Qy	467	ACCAGGCGAGTACCACTGTGGCCTGG	493	
Db	983	ACGCGGCGGAGTACACGTGTGAGTGG	1009	

RESULT 8	
ABZ224581	ABZ224581 standard; cDNA; 7564 BP.
XX	
XX	ABZ224581;
XX	
XX	31-MAR-2003 (first entry)
XX	
DE	Human cell adhesion and extracellular matrix protein 4 cDNA.
XX	
KW	Cell adhesion and extracellular matrix protein 4; CADECM-4; human;
KW	anti-HIV; virucide; antiallergic; antiinflammatory; antihaemic;
KW	antiparkinsonian; nootropic; anticonvulsant; antifertility;
KW	antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;
KW	cytostatic; hepatotropic; dermatological; antidiabetic; nephrotropic;
KW	antigout; thyromimetic; neuroprotective; osteopathic; aniarthritic;
KW	antiparasitic; anhelminthic; antipneumatic; antipsoasitic; fungicide;
KW	antirheumatic; haemostatic; antibacterial; protozoaside; gene; ss.
XX	gynaecological; titin; gene therapy; Gene; ss.
XX	
OS	Homo sapiens.

XX	WO20028322-A2.
PN	
XX	
XX	07-NOV-2002.
XX	
XX	01-MAY-2002; 2002WO-US013874.
XX	
PR	02-MAY-2001; 2001US-0288290P.
PR	21-MAY-2001; 2001US-0292468P.
PR	15-JUN-2001; 2001US-0298616P.
PR	28-JUN-2001; 2001US-0301672P.
PR	04-JAN-2002; 2002US-0345009P.
PR	

XX (INCY-) INCYTE GENOMICS INC.
XX
XX
PI Yee H, Lee EA, Duggan BM, Honchell CD, Ding L,
PI Hillman JF, Baughn MR, Kallick DA, Lee S, Warren BA, Xu Y, Tran UK,
PI Lal PG, Thornton M, Hafalia AJA, Yao MG, Nguyen DB, Gandhi AR;
PI Khan FA, Walhia NK, Griffin JA, Chinn AM, Elliott VS, Ramkumar J;
PI Arvizu CS, Forsythe JG,
XX
XX WPI: 2003-167112/16.
DR P-FSDB: ABF58227.
DR
XX
XX New human cell adhesion and extracellular matrix proteins, useful for
PT diagnosing, treating or preventing autoimmune or inflammatory disorder
PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
PT cancer or hepatitis.
XX
XX Claim 12: Page 168-170: 178pp: English.

The present sequence is that of Incyte clone 7326129CB1 cDNA encoding human cell adhesion and extracellular matrix protein 4 (CADEM-4). A representative cDNA library for the full-length polynucleotide is MUSE129D02, constructed from the muscle tissue RNA of a Caucasian adult man. Homology data suggest the encoded protein to a titin muscle protein. The invention provides CADEM-1 to -11 polypeptides (see ABP58224-34) and polynucleotides (see ABP24578-88), expression vectors, host cells, antibodies, agonists and antagonists. These are useful for diagnosing, treating or preventing disorders associated with aberrant expression of CADEM, particularly cell proliferative disorders (e.g. arteriosclerosis, atherosclerosis, cirrhosis, hepatitis, psoriasis, paroxysmal nocturnal haemoglobinuria, polycythemia vera, porphyria, primary thrombocytopaenia or cancer), developmental disorders (e.g. renal tubular acidosis, anaemia or mental retardation), neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or epilepsy), reproductive disorders (e.g. infertility or a disruption in the menstrual cycle), or autoimmune/inflammatory disorders (e.g. AIDS, allergy, asthma, autoimmune thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral, bacterial, fungal, parasitic, protozoal or helminthic infections.

Query Match	4.9%	Score 53.4	DB 8	Length 7564
Best Local Similarity	53.6%	Pred. No. 0.065		
Matches 111	Conservative 0	Mismatches 96	Indels 0	Gaps 0
QY	287	CCCTGAGAGCCGGCAGGTGGCAGCTGGTCAAGATGTCTCTGTGAGCTTTGAGGTGGTGG	346	
Db	2612	CCCTGGAAGACCACTGGTGGCCCGGAGGAGACGTGGAGTTCGCTGTGAGCTGTCCAC	2671	
QY	347	CTGAGGCTGGTGAAGTCAATCTGGCCACAAGGAAATGGAGCGCATCCAGCCCGGTGGGGCGT	406	
Db	2672	GGCGCGGAAACGCCCGGTGCACCTGGCTGAAGACAGGAAGGCGCATCCGCAAGAGCCAGAAGT	2731	
QY	407	TCGAGTGTGTCTCCAGGGTGGCCACACAGATGCTGGTGATCAAGGGCTTTCACGCGCAGAAG	466	
Db	2732	ATGATGTGGTCTCGGAGGGACACATGCCCATGCTGGTTCATCCGCGGGGCGCTCGCTCAAGG	2791	
QY	467	ACCAGGCGCAGTATACCAGTGTGGCCTGG	493	
Db	2792	ACGCGGGCAGGTACAGTGTGAGGTGG	2818	

RESULT 9
ABN21414
ID ABN21414 standard; cDNA; 1005 BP.
XX
AC ABN21414;
XX
DT 24-JUN-2002 (first entry)

XX DE Human ORFX polynucleotide sequence SEQ ID NO:11305.

XX DE Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;

KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;

KW degenerative disorder; osteoarthritis; neurodegenerative disorder;

KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;

KW hypertension; hypothyroidism; cholesterol ester storage disease;

KW immune deficiency; immune disorder; infectious disease;

KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;

KW myasthenia gravis; gene; ss.

OS Homo sapiens.

XX WO200192523-A2.

XX PD 06-DEC-2001.

XX PF 29-MAY-2001; 2001WO-US010836.

XX PF 30-MAY-2000; 2000US-0206132P.

PR 29-AUG-2000; 2000US-0228716P.

XX PA (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach MD;

XX PI WPI; 2002-106308/14.

DR P-PSDB; ABP05662.

XX PT Novel human polypeptides and polynucleotides useful for diagnosing,

PT preventing and treating cardiovascular disease, neurodegenerative,

PT hyperproliferative disorders and autoimmune disorders.

XX PS Disclosure; SEQ ID NO 11305; 1037pp; English.

XX CC The present invention describes substantially purified human proteins

CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1

CC in the specification). ABN15762 to ABN27252 encode the human ORFX

CC proteins given in ABP00010 to ABP11500. ORFX proteins are useful for

CC treating or preventing a pathology associated with an ORFX-associated

CC disorder in humans, and in the manufacture of a medicament for treating a

CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide

CC sequences can be used in gene therapy. ORFX sequences can be used in the

CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,

CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,

CC osteoarthritis, neurodegenerative disorders, disorders related to organ

CC transplantation, cardiovascular diseases, diabetes mellitus, systemic

CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester

CC storage disease, various immune deficiencies and disorders, infectious

CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid

CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host

CC disease and autoimmune inflammatory eye disease. ORFX proteins are also

CC useful for treating burns, incisions, ulcers, for treating osteoporosis,

CC bone degenerative disorders, or periodontal disease, and for gut

CC protection or regeneration and treatment of lung or liver fibrosis,

CC reperfusion injury in various tissues and conditions resulting from

CC systemic cytokine damage. N.B. The sequence data for this patent did not

CC form part of the printed specification, but was obtained in electronic

CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 1005 BP; 209 A; 303 G; 340 G; 152 T; 0 U; 1 Other;

Query Match 4.6%; Score 50.2; DB 6; Length 1005;

Best Local Similarity 52.7%; Pred. No. 0.23;

Matches 109; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Oy 287 CCCCTGAGAGCCGGCAGTGCGTGTGAGATGTCCTCTGAGAGTTGAGGTGGG 346

DB 353 CCCTGGAAGACCAAGTGGTGGCGCCAGGGGAGGACGTGGAGCTGTGAGCTGTAC 412

Oy 347 CTGAGGCTGTGAGGTGTCATCTGGCACAGGGNATGGAGGCATCCGCCCGTGGCGGT 406

DB 412

XX DB 413 GGGCGGGAACGCCCGTGCACCTGGCTGAAGGACGGAGGCCCATCCGCAAGGCCAGAAAGT 472

Oy 407 TCGAGGTGGTCTCCACAGGTCGGCAACAGATGCTGGTGTATCAAGGGCTTCACGGCAGAAG 466

DB 473 ATGATGTGGTCTCGAGGGGACGATGGCCATGCTGGTCAATCCGCGGGGCTCGCTCAAGG 532

Oy 467 ACCAGGCGAGTACCACTGTGGCCCTGG 493

DB 533 ACGCGGGCGAGTACACGTCGTGAGGTGG 559

RESULT 10

ADA71938/c

ID ADA71938 standard; DNA; 2000 BP.

XX AC ADA71938;

XX DT 20-NOV-2003 (first entry)

XX DE Rice gene, SEQ ID 5263.

XX KW Plant; bacterial infection; fungal infection; viral infection; rice;

XX KW gene; ds.

XX OS Oryza sativa.

XX PN WO2003000898-A1.

XX PD 03-JAN-2003.

XX PF 22-JUN-2001; 2001WO-IB001105.

XX PR 22-JUN-2001; 2001WO-IB001105.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;

PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;

XX WPI; 2003-175290/17.

XX PT Identifying at least one gene involved in plant resistance or response to

PT pathogenic infection for conferring resistance or tolerance to a plant to

PT bacterial, fungal or viral infection by determining or detecting plant

XX gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.

XX CC The present invention relates to a method (M1) for identifying genes

CC involved in plant resistance or response to pathogenic infection. M1

CC comprises identifying a gene whose expression is significantly altered in

CC the incompatible interaction of plant gene expression relative to

CC expression of the gene in an uninfected plant, in a mutant plant that

CC does not express a gene associated with response to pathogenic infection,

CC or in a corresponding incompatible or compatible interaction. (M1) is

CC useful for conferring resistance to resistance or tolerance to a plant to

CC bacterial, fungal or viral infection. The present sequence was used to

XX illustrate the invention.

SQ Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

Query Match 4.5%; Score 49.4; DB 8; Length 2000;

Best Local Similarity 8.8%; Pred. No. 0.39;

Matches 59; Conservative 306; Mismatches 302; Indels 2; Gaps 1;

Oy 263 AGCTGTGCTCCCGCCCTCAGCCCTCAGCCCTGAGAGCCGCGAGTGGCAGCTGGTGAAGATG 322

DB 705 RSYRRRRWYWKRWYRYRYRWYRWSRTRMRMRKRKRWAGASMKSCMYRWGARSMMYS 646

Oy 323 TCTCTGGAGCTTGAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 382

DB 645 KYSCAKCKKTRWTSSYMTSGMYSSYKMSWTSKMSYMGKWTCTMTYMKGSTR 586

Qy	383	ACGCATCCAGCCCGGTGGCCGGTTCGAGGTGGTCTCCAGGGTCGGCAACAGATGCTGG	442
Db	585	SKGRWSGMSRMYMRWKRKYMRWYKWKCTWERQCYRWGYTWYTTSSRWMYTGR	526
Qy	443	TGATCAAGGGCTTCACGGCAGAAGACACAGGGCGAGTACCACCTGTGGCTGGCTCAGGGCT	502
Db	525	YKARTSRKRYWYKYRKYCWYYYGYMYMKCSYMMRYGYCKACKCKCYAMCAKAAYSGMV	466
Qy	503	CCATCTGCCCTCGGCTGCACCTCCAGCTGGCACTGAGCCACAGCTCTGTGGATGAGG	562
Db	465	MYWRKYSKNWRMSTKYMWSMYKKRSMKYGAKGCGKMWYCSYGMKWTYMGSYK	406
Qy	563	C--CCCTCAGCCAGCTTGCCCCCGAGCAGCAGGAGGTGACCTGACCTACTGTG	620
Db	405	YSRCYKMYRMWYKGMWYMYYSAYSSMTWYTYAKYWKYKRRGTMSWYKSYKKY	346
Qy	621	GGAGCCCTGGCTCGGAACGTCGCATGAGCGGTGAGCCACGCTGGACTCCATTAGCGA	680
Db	345	CTWVCYMKCMRYRWRKMRKTKYSKRYCWRYATCYWCCCYRKRGMYSRRSMRTAGK	286
Qy	681	GCTGCCAGGAGGACGCCCGCTCGCAGCGCTGCCACAGGAGCAGAGGAGTGGCACC	740
Db	285	WKMRWSRWSRYSYWKYKWKWKYSYMSYGWARSSTWRSAAKRTYKGYSTSRKM	226
Qy	741	TGATCTCTCAAGCTACTCCACGGCCGATCAGCTGGCCGCACTGGAGATGCTGACCT	800
Db	225	MRACBMTYSACRYRSTYSYCGSYGSSKWKYMSKSCSMRMTCSWCCYTCYYGAMCW	166
Qy	801	CTCACACACCACTGTGATGATGATGATCCCGGCGAGCACCCCTTCCCTGGTCACTACCT	860
Db	165	SCCMSMYMGSCGCVTRGWKRSKYSMCCKKYCSCTKYCSYTGYYRYCKWKYSYKYCY	106
Qy	861	CAAGAAGCTGGGAGGCAGCAGCACTCACCACTGGCCAGCAAGGTTGGGGGCCCGCAGC	920
Db	105	YCYCYWYSYMRMYMKCRSCRSWSWSCAYCSTSTSRWMSMYAAKMGVCGSSGYMR	46
Qy	921	CCCCTCTGT	929
Db	45	SKSCWYSK	37
RESULT 11			
ADQ22659			
ID	ADQ22659	standard; DNA; 1813 BP.	
XX	ADQ22659;		
XX	26-AUG-2004	(first entry)	
XX	Human soft tissue sarcoma-upregulated DNA - SEQ ID 5479.		

PT sarcoma.

PS Example 2; SEQ ID NO 5479; 210pp; English.

XX

CC The invention relates to a novel method for detecting soft tissue sarcoma which comprises obtaining a first soft tissue sample from an individual and a normal soft tissue sample from the same or different individual, determining the expression of a gene in both samples and comparing the expression of the gene in both soft tissue samples, where a higher level of protein expression in the first soft tissue sample indicates the presence of soft tissue sarcoma. The method of the invention has cyrostatic applications and may be useful for detecting soft tissue sarcoma, possibly via gene therapy or vaccine production. The nucleic acid sequences may be useful in diagnostic and screening applications. The current sequence is that of a human soft tissue sarcoma-upregulated DNA of the invention. The current sequence is not shown within the specification per se but was submitted in CD format by the inventor.

XX Sequence 1813 BP; 410 A; 590 C; 505 G; 290 T; 0 U; 18 Other:

Early detection of soft tissue sarcoma comprises determining expression of a gene in a first soft tissue sample and a normal soft tissue sample and comparing the gene expression, also useful in treating soft tissue


```

Db      544 ATCCACATCCGAGAG 558      || | | | ||||
RESULT 14
ADQ21633
ID      ADQ21633 standard; DNA; 3613 BP.
XX
XX      AC      ADQ21633;
XX
XX      DT      26-AUG-2004 (first entry)
XX
DE      Human soft tissue sarcoma-upregulated DNA - SEQ ID 4453.
XX
XX      KW      soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX      KW      ds.
XX
XX      OS      Homo sapiens.
XX
XX      FN      WO2004048938-A2.
XX
XX      PD      10-JUN-2004.
XX
XX      PF      26-NOV-2003; 2003WO-US038193.
XX
XX      PR      26-NOV-2002; 2002US-0429739P.
XX
XX      PA      (PROT-) PROTEIN DESIGN LABS INC.
XX
XX      PI      Aziz N, Ginsburg WM, Zlotnik A;
XX
XX      DR      WPI; 2004-441208/41.
XX
XX      PT      Early detection of soft tissue sarcoma comprises determining expression
XX      PT      of a gene in a first soft tissue sample and a normal soft tissue sample
XX      PT      and comparing the gene expression, also useful in treating soft tissue
XX      PT      sarcoma.
XX
XX      PS      Example 2; SEQ ID NO 4453; 210pp; English.
XX
XX      CC      The invention relates to a novel method for detecting soft tissue sarcoma
XX      CC      which comprises obtaining a first soft tissue sample from an individual
XX      CC      and a normal soft tissue sample from the same or different individual,
XX      CC      determining the expression of a gene in both samples and comparing the
XX      CC      expression of the gene in both soft tissue samples, where a higher level
XX      CC      of protein expression in the first soft tissue sample indicates the
XX      CC      presence of soft tissue sarcoma. The method of the invention has
XX      CC      cytostatic applications and may be useful for detecting soft tissue
XX      CC      sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX      CC      acid sequences may be useful in diagnostic and screening applications.
XX      CC      The current sequence is that of a human soft tissue sarcoma-upregulated
XX      CC      DNA of the invention. The current sequence is not shown within the
XX      CC      specification per se but was submitted in CD format by the inventor.
XX
XX      SQ      Sequence 3613 BP; 778 A; 1026 C; 1185 G; 624 T; 0 U; 0 Other;

Query Match      4.2%; Score 46.6; DB 12; Length 3613;
Best Local Similarity 53.8%; Pred. No. 1.8;
Matches 119; Conservative 0; Mismatches 99; Indels 3; Gaps 1;

Qy      6 CAGCCACACTGTGACACTGTTCTGGGAGCTCCCATGAGTGATGAGAGCGGTGTCCTG 65
Db      1743 CAGCCAGGGCATCATCTGACATGGACAGCACCTCGGGGCCCGGAGCGCCACATCCT 1802

Qy      66 TGGCTACCGCTGGAGTG---AAGGAGGGGGCCACAGGCCAGTGGCGGCTGTGCCACGA 122
Db      1803 GGGCTACCTGATCGAGAGGCGTAAAGAGGGGAGCAACACTGGACGCGAGTGAACGACCA 1862

Qy      123 GCTGGTCCCTGGAGCCCGAGTGTGTGGGAGTGCTGGCCCGCGGGAGAGCCTTACCCTTT 182
Db      1863 GCCGGTCCCTGAGAGGAGTGAGCGTGGCGGAGCTGCGGAGGCGCTGTCTAGTATGAGTT 1922

Qy      183 CCGTGTGGCAGCTGTGGGCCCTGTGGGTGTGGGTGTGGGGAACCGG 223

Db      1923 CCGGGTCACAGCTGTGGCTCCCTCAGGTCCGGGAGGCGCTG 1963

RESULT 15
ADQ25448
ID      ADQ25448 standard; DNA; 4176 BP.
XX
XX      AC      ADQ25448;
XX
XX      DT      26-AUG-2004 (first entry)
XX
DE      Human soft tissue sarcoma-upregulated DNA - SEQ ID 8268.
XX
XX      KW      soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX      KW      ds.
XX
XX      OS      Homo sapiens.
XX
XX      FN      WO2004048938-A2.
XX
XX      PD      10-JUN-2004.
XX
XX      PF      26-NOV-2003; 2003WO-US038193.
XX
XX      PR      26-NOV-2002; 2002US-0429739P.
XX
XX      PA      (PROT-) PROTEIN DESIGN LABS INC.
XX
XX      PI      Aziz N, Ginsburg WM, Zlotnik A;
XX
XX      DR      WPI; 2004-441208/41.
XX
XX      PT      Early detection of soft tissue sarcoma comprises determining expression
XX      PT      of a gene in a first soft tissue sample and a normal soft tissue sample
XX      PT      and comparing the gene expression, also useful in treating soft tissue
XX      PT      sarcoma.
XX
XX      PS      Example 2; SEQ ID NO 8268; 210pp; English.
XX
XX      CC      The invention relates to a novel method for detecting soft tissue sarcoma
XX      CC      which comprises obtaining a first soft tissue sample from an individual
XX      CC      and a normal soft tissue sample from the same or different individual,
XX      CC      determining the expression of a gene in both samples and comparing the
XX      CC      expression of the gene in both soft tissue samples, where a higher level
XX      CC      of protein expression in the first soft tissue sample indicates the
XX      CC      presence of soft tissue sarcoma. The method of the invention has
XX      CC      cytostatic applications and may be useful for detecting soft tissue
XX      CC      sarcoma, possibly via gene therapy or vaccine production. The nucleic
XX      CC      acid sequences may be useful in diagnostic and screening applications.
XX      CC      The current sequence is that of a human soft tissue sarcoma-upregulated
XX      CC      DNA of the invention. The current sequence is not shown within the
XX      CC      specification per se but was submitted in CD format by the inventor.
XX
XX      SQ      Sequence 4176 BP; 967 A; 1133 C; 1364 G; 705 T; 0 U; 7 Other;

Query Match      4.2%; Score 46.6; DB 12; Length 4176;
Best Local Similarity 53.8%; Pred. No. 1.8;
Matches 119; Conservative 0; Mismatches 99; Indels 3; Gaps 1;

Qy      6 CAGCCACACTGTGACACTGTTCTGGGAGCTCCCATGAGTGATGAGAGCGGTGTCCTG 65
Db      1745 CAGCCAGGGCATCATCTGACATGGACAGCACCTCGGGGCCCGGAGCGCCACATCCT 1804

Qy      66 TGGCTACCGCTGGAGTG---AAGGAGGGGGCCACAGGCCAGTGGCGGCTGTGCCACGA 122
Db      1805 GGGCTACCTGATCGAGAGGCGTAAAGAGGGGAGCAACACTGGACGCGAGTGAACGACCA 1864

Qy      123 GCTGGTCCCTGGAGCCCGAGTGTGTGGGAGTGCTGGCCCGCGGGAGAGCCTTACCCTTT 182
Db      1865 GCCGGTCCCTGAGAGGAGTGAGCGTGGCGGAGCTGCGGAGGCGCTGTCTAGTATGAGTT 1924

Qy      183 CCGTGTGGCAGCTGTGGGCCCTGTGGGTGTGGGTGTGGGGAACCGG 223
```


Db 1925 CCGGTCACAGCTGTGGCTCCCTCAGGTCCCGGAGGCTG 1965

Search completed: March 20, 2005, 18:55:07
Job time : 644.09 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:15:10 ; Search time 232.766 Seconds

(without alignments)

10198.307 Million cell updates/sec

Title: US-10-077-130-6_COPY_8500_8900

Perfect score: 401

Sequence: 1 gaagaccagtggtggcgcc.....caccagccagccagggagg 401

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: geneseqn1980s:*

2: geneseqn1990s:*

3: geneseqn2000s:*

4: geneseqn2001as:*

5: geneseqn2001bs:*

6: geneseqn2002as:*

7: geneseqn2002bs:*

8: geneseqn2003as:*

9: geneseqn2003bs:*

10: geneseqn2003cs:*

11: geneseqn2003ds:*

12: geneseqn2004as:*

13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	401	100.0	2534	11 ADM03261	Adm03261 Human cDN
2	401	100.0	7564	8 ABZ24581	Abz24581 Human cel
3	401	100.0	14061	6 ABV99363	Abv99363 Human NOV
4	401	100.0	14109	6 ABV99362	Abv99362 Human NOV
5	401	100.0	20489	12 ADQ922881	Adq22881 Human sof
6	401	100.0	24120	8 ABX11642	Abx11642 Human ser
7	393	98.0	1005	6 ABN21414	Abn21414 Human ORF
8	60.6	15.1	2559	8 ACC46215	Acc46215 Human dit
9	60.6	15.1	2737	4 AAI59251	Aai59251 Human pol
10	60.6	15.1	3999	4 AAI161037	Aai161037 Human pol
11	60.6	15.1	5378	13 ADR67197	Adr67197 Human bla
12	60.6	15.1	5382	10 ADD14722	Add14722 Human arc
13	60.6	15.1	20565	4 AAK89083	Aak89083 Human dig
14	60.6	15.1	20565	4 AAS28908	Aas28908 Human lmm
15	60.6	15.1	20565	4 AAL03357	Aal03357 Human rep
16	60.6	15.1	20565	10 ADB31749	Adb31749 Human nov
17	59.8	14.9	31595	10 ADF81661	Adf81661 Leukaemia
18	59.8	14.9	81940	4 AAS05390	Aas05390 Human tit
19	59.8	14.9	81940	6 ABK64829	Abk64829 Human ben
20	59.8	14.9	81940	12 ADQ17315	Adq17315 Human sof

21	59.8	14.9	93801	9	ABX13540	
22	59.8	14.9	103052	13	ADQ89963	
23	59	14.7	2184	4	AAH18256	
24	56	14.0	716	6	ABK35726	
25	56	14.0	1266	6	ABK35725	
26	56	14.0	1645	5	AAF24162	
27	56	14.0	1948	2	ABV74346	
28	56	14.0	1949	2	ABV74347	
29	56	14.0	2170	6	ABQ54970	
30	56	14.0	9591	4	AAS28906	
31	56	14.0	9591	4	AAL03355	
32	56	14.0	9591	10	ADB31747	
33	56	14.0	12415	4	AAS28907	
34	56	14.0	12415	4	AAL03356	
35	56	14.0	12415	10	ADB31748	
C	36	55.8	13.9	466	12	ACH89436
37	55	13.7	724	4	AAH08421	
38	54.8	13.7	4302	13	ACN41312	
39	54.8	13.7	7328	4	AAK89081	
40	54.4	13.6	2768	8	ABX71198	
C	41	52.8	13.2	3956	13	ADR08183
42	52.6	13.1	572	5	AAS42508	
43	52.6	13.1	1251	10	ADC30509	
44	52.6	13.1	2488	4	AAH18453	
45	51.2	12.8	2000	8	ADA71938	

ALIGNMENTS

RESULT 1	
ADM03261	
ID	ADM03261 standard; cDNA; 2534 BP.
XX	
AC	ADM03261;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Human cDNA of the invention SEQ ID NO:1946.
XX	
KW	ss; Gene; human; Gene therapy; diagnostic marker; pharmaceutical.
XX	
OS	Homo sapiens.
XX	
PN	EPI347046-A1.
XX	
PD	24-SEP-2003.
XX	
PF	12-APR-2002; 2002EP-00008400.
XX	
PR	22-MAR-2002; 2002JP-00137785.
XX	
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.
XX	
PI	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI	Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
XX	Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
DR	WPI; 2003-723558/69.
DR	P-PsDB; ADM05704.
XX	
PT	New polynucleotides and polypeptides are useful in gene therapy, for
PT	developing a diagnostic marker or medicines for regulating their
PT	expression and activity, or as a target of gene therapy.
XX	
PS	Claim 1; SEQ ID NO 1946; 305pp; English.
XX	
CC	The invention relates to a novel human polynucleotide and the encoded
CC	polypeptide. A polynucleotide of the invention may have a use in gene
CC	therapy. An oligonucleotide of the invention ADM06202-ADM06773 is useful
CC	as a primer for synthesizing the polynucleotide or as a probe for
CC	detecting the polynucleotide. The polynucleotides ADM01316-ADM03758 are
CC	useful in gene therapy, for developing a diagnostic marker or medicines

CC for regulating their expression and activity, or as a target of gene
CC therapy. The proteins ADM03759-ADM06201 encoded by the polynucleotides
CC are useful as pharmaceutical agents. The present sequence represents a
CC cDNA sequence of the invention.

XX SQ Sequence 2534 BP; 551 A; 679 C; 778 G; 526 T; 0 U; 0 Other;

Query Match 100.0%; Score 401; DB 11; Length 2534;
Best Local Similarity 100.0%; Pred. No. 1.4e-86;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAAGACAGTGGTGGCGCCAGGGAGGAGCTGGCTGTGAGTGTGACGGCG 60
Db |||||
808 GAAGACAGTGGTGGCGCCAGGGAGGAGCTGGCTGTGAGTGTGACGGCG 867
QY 61 GGAACGCCCTGGCTGGCTGAAGGACAGGAGGCCATCCGCAAGACCGCAAGTATGAT 120
Db |||||
868 GGAACGCCCTGGCTGGCTGAAGGACAGGAGGCCATCCGCAAGACCGCAAGTATGAT 927
QY 121 GTGGTCTGGAGGCGACGATGGCCATGTGTCATCCGCGGGGCTCGCTCAAGGACGCG 180
Db |||||
928 GTGGTCTGGAGGCGACGATGGCCATGTGTCATCCGCGGGGCTCGCTCAAGGACGCG 987
QY 181 GCGGAGTACACGTGTGAGTGGAGGCTTCCAGAGCACAGCCAGCTCCATGTGAAGAA 240
Db |||||
988 GCGGAGTACACGTGTGAGTGGAGGCTTCCAGAGCACAGCCAGCTCCATGTGAAGAA 1047
QY 241 AAAGCAAATCTGTTTACAGAGGAGTGTACCAATCTGCAAGTGGAGGAGAAAGGCACGCT 300
Db |||||
1048 AAAGCAAATCTGTTTACAGAGGAGTGTACCAATCTGCAAGTGGAGGAGAAAGGCACGCT 1107
QY 301 GTGTTACGTGCAAGCGAGGACACCCGCGGCCACAGTGAAGTGGCGCAAGGGCTCTTG 360
Db |||||
1108 GTGTTACGTGCAAGCGAGGACACCCGCGGCCACAGTGAAGTGGCGCAAGGGCTCTTG 1167
QY 361 GAGCTACGGGCTCGAGGAGGAGCAGCCAGCCAGCCAGGAGGG 401
Db |||||
1168 GAGCTACGGGCTCGAGGAGGAGCAGCCAGCCAGGAGGG 1208

RESULT 2

ID ABZ24581
ABZ24581 standard; cDNA; 7564 BP.

AC ABZ24581;

XX 31-MAR-2003 (first entry)

DE Human cell adhesion and extracellular matrix protein 4 cDNA.

XX Cell adhesion and extracellular matrix protein 4; CADECM-4; human;
KW anti-HIV; virucide; antifungal; anti-inflammatory; antianemia;
KW antiparkinsonian; neurotropic; anticonvulsant; antifertility;
KW antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;
KW cytosolic; hepatotropic; dermatological; antidiabetic; nephrotropic;
KW angiotensin; neurotrophic; neuroprotective; osteopathic; antithyroid;
KW antiparasitic; antihelminthic; antiparasitic; uropathic; ophthalmological;
KW antirheumatic; haemostatic; antibacterial; protozoacide; fungicide;
KW synaenological; titin; gene therapy; gene; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 241..7227
FT /*tag= a
FT /product= "Human CADECM-4"

XX WO20028322-A2.

XX 07-NOV-2002.

XX 01-MAY-2002; 2002WO-US013874.

PR 02-MAY-2001; 2001US-0298290P.
PR 21-MAY-2001; 2001US-0292468P.
PR 15-JUN-2001; 2001US-0298616P.
PR 28-JUN-2001; 2001US-0301672P.
PR 04-JAN-2002; 2002US-0345008P.
XX (INCY-) INCYTE GENOMICS INC.
XX Yue H, Lee EA, Duggan BM, Thangavelu K, Honchell CD, Ding L;
PI Hillman JL, Baughn MR, Kallick DA, Lee S, Warren BA, Xu Y, Tran UK;
PI Lal PG, Thornton M, Hafalia AJA, Yao MG, Nguyen DB, Gandhi AR;
PI Khan FA, Walha NK, Griffin JA, Chinn AM, Elliott VS, Ramkumar J;
XX Arvizu CS, Forsythe LJ;
DR WPI; 2003-167112/16.
DR P-PSDB; ABP58227.
XX New human cell adhesion and extracellular matrix proteins, useful for
diagnosing, treating or preventing autoimmune or inflammatory disorder
PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
PT cancer or hepatitis.
XX

PS Claim 12; Page 168-170; 178pp; English.

XX The present sequence is that of Incyte clone 7326129CB1 cDNA encoding
human cell adhesion and extracellular matrix protein 4 (CADECM-4). A
representative cDNA library for the full-length polynucleotide is
MUSLDR02, constructed from the muscle tissue RNA of a Caucasian adult
man. Homology data suggest the encoded protein to a titin muscle protein.
The invention provides CADECM-1 to -11 polypeptides (see ABP58224-34) and
polynucleotides (see ABZ24578-88), expression vectors, host cells,
antibodies, agonists and antagonists. These are useful for diagnosing,
treating or preventing disorders associated with aberrant expression of
CADECM, particularly cell proliferative disorders (e.g. arteriosclerosis,
atherosclerosis, cirrhosis, hepatitis, psoriasis, paroxysmal nocturnal
haemoglobinuria, polycythaemia vera, psoriasis, primary
thrombocytopenia or cancer), developmental disorders (e.g. renal
tubular acidosis, anaemia or mental retardation), neurological disorders
(e.g. Alzheimer's disease, Parkinson's disease or epilepsy), reproductive
disorders (e.g. infertility or a disruption in the menstrual cycle), or
autoimmune/inflammatory disorders (e.g. AIDS, allergy, asthma, autoimmune
thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease,
Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis,
osteoarthritis, osteoporosis, pancreatitis, Reiter's syndrome, rheumatoid
arthritis, Sjogren's syndrome, uveitis), or viral, bacterial, fungal,
parasitic, protozoal or helminthic infections

XX SQ Sequence 7564 BP; 1594 A; 1999 C; 2711 G; 1260 T; 0 U; 0 Other;

Query Match 100.0%; Score 401; DB 8; Length 7564;
Best Local Similarity 100.0%; Pred. No. 1.7e-86;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GAAGACAGTGGTGGCGCCAGGGAGGAGCTGGCTGTGAGTGTGACGGCG 60
Db |||||
2617 GAAGACAGTGGTGGCGCCAGGGAGGAGCTGGCTGTGAGTGTGACGGCG 2676
QY 61 GGAACGCCCTGGCTGGCTGAAGGACAGGAGGCCATCCGCAAGACCGCAAGTATGAT 120
Db |||||
2677 GGAACGCCCTGGCTGGCTGAAGGACAGGAGGCCATCCGCAAGACCGCAAGTATGAT 2736
QY 121 GTGGTCTGGAGGCGACGATGGCCATGTGGTGTATCCCGGGGCTCGCTCAAGGACGCG 180
Db |||||
2737 GTGGTCTGGAGGCGACGATGGCCATGTGGTGTATCCCGGGGCTCGCTCAAGGACGCG 2796
QY 181 GCGGAGTACAGTGTGAGGTGGAGGCTTCCAAAGACACAGCCAGCTCCATGTGAAGAA 240
Db |||||
2797 GCGGAGTACAGTGTGAGGTGGAGGCTTCCAAAGACACAGCCAGCTCCATGTGAAGAA 2856
QY 241 AAAGCAAATCTGTTTACAGAGGAGTGTACCAATCTGCAAGTGGAGGAGAAAGGCACGCT 300
Db |||||
2857 AAAGCAAATCTGTTTACAGAGGAGTGTACCAATCTGCAAGTGGAGGAGAAAGGCACGCT 2916

QY 301 GTGTTACGTCACAGCAGGACCCCGGGCCACAGTACCTGCGGCAAGGGCTCTTG 360
 DB 2917 GTGTTACGTCACAGCAGGACCCCGGGCCACAGTACCTGCGGCAAGGGCTCTTG 2976
 QY 361 GAGCTACGGGCTCAGGGAAGCACCAGCCAGCCAGCCAGGAGGG 401
 DB 2977 GAGCTACGGGCTCAGGGAAGCACCAGCCAGCCAGGAGGG 3017

RESULT 3

ABV99363
 ID ABV99363 standard; DNA; 14061 BP.

XX
 AC ABV99363;
 XX

DT 27-JAN-2003 (first entry)
 XX

DE Human NOVA3b coding sequence.
 XX

KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
 KW antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;
 KW nootropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
 KW antifertility; cerebroprotective; gene therapy; NOVA; NOV; fertility;
 KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
 KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
 KW immune disorder; haematopoietic disorder; cardiovascular disorder;
 KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
 KW metabolic syndrome X; wasting disorder; cell differentiation; gene;
 KW cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.

XX Homo sapiens.
 OS

XX WO200272771-A2.
 PN

XX 19-SEP-2002.
 PD

XX 08-MAR-2002; 2002WO-US007288.
 PF

PR 08-MAR-2001; 2001US-0274101P.
 PR 08-MAR-2001; 2001US-0274194P.
 PR 08-MAR-2001; 2001US-0274281P.
 PR 08-MAR-2001; 2001US-0274322P.
 PR 09-MAR-2001; 2001US-0274849P.
 PR 12-MAR-2001; 2001US-0275233P.
 PR 13-MAR-2001; 2001US-0275578P.
 PR 13-MAR-2001; 2001US-0275579P.
 PR 13-MAR-2001; 2001US-0275601P.
 PR 14-MAR-2001; 2001US-0276000P.
 PR 16-MAR-2001; 2001US-0276776P.
 PR 19-MAR-2001; 2001US-0276994P.
 PR 20-MAR-2001; 2001US-0277239P.
 PR 20-MAR-2001; 2001US-0277321P.
 PR 20-MAR-2001; 2001US-0277327P.
 PR 20-MAR-2001; 2001US-0277338P.
 PR 21-MAR-2001; 2001US-0277791P.
 PR 23-MAR-2001; 2001US-0277833P.
 PR 26-MAR-2001; 2001US-0278894P.
 PR 27-MAR-2001; 2001US-0278999P.
 PR 27-MAR-2001; 2001US-0279038P.
 PR 28-MAR-2001; 2001US-0279344P.
 PR 30-MAR-2001; 2001US-0279995P.
 PR 30-MAR-2001; 2001US-0280233P.
 PR 02-APR-2001; 2001US-0280802P.
 PR 02-APR-2001; 2001US-0280822P.
 PR 04-APR-2001; 2001US-0280900P.
 PR 13-APR-2001; 2001US-0281194P.
 PR 30-APR-2001; 2001US-0287424P.
 PR 03-MAY-2001; 2001US-0288066P.
 PR 03-MAY-2001; 2001US-0288342P.
 PR 03-MAY-2001; 2001US-0288528P.

PR 15-MAY-2001; 2001US-0291190P.
 PR 16-MAY-2001; 2001US-0291099P.
 PR 30-MAY-2001; 2001US-0291240P.
 PR 31-MAY-2001; 2001US-0294485P.
 PR 31-MAY-2001; 2001US-0294899P.
 PR 18-JUN-2001; 2001US-0299027P.
 PR 19-JUN-2001; 2001US-0299303P.
 PR 19-JUN-2001; 2001US-0299310P.
 PR 10-JUL-2001; 2001US-0304354P.
 PR 31-JUL-2001; 2001US-0309198P.
 PR 16-AUG-2001; 2001US-0312903P.
 PR 10-SEP-2001; 2001US-0318462P.
 PR 12-SEP-2001; 2001US-0318770P.
 PR 27-SEP-2001; 2001US-0325430P.
 PR 27-SEP-2001; 2001US-0325681P.
 PR 18-OCT-2001; 2001US-0330380P.
 PR 31-OCT-2001; 2001US-0335301P.
 PR 14-NOV-2001; 2001US-0332172P.
 PR 14-NOV-2001; 2001US-0332271P.
 PR 14-NOV-2001; 2001US-0332272P.
 PR 14-NOV-2001; 2001US-0333184P.
 PR 21-NOV-2001; 2001US-0333272P.
 PR 03-DEC-2001; 2001US-0332094P.
 PR 03-DEC-2001; 2001US-0337436P.
 PR 04-DEC-2001; 2001US-0338092P.
 PR 03-JAN-2002; 2002US-0345705P.
 PR 08-MAR-2002; 2002US-00093463.
 XX

(CURA-) CURAGEN CORP.

PA Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
 XX Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CAM;
 PI Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
 PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
 PI Taupier RJ, Padigaru M, Shenoy SG, Kekuda R, Gusev VY, Pochart PF;
 PI Zhong M;
 XX

WPI; 2002-732824/79.
 P-PSDB; ABP70085.

New NOVA polypeptides and polynucleotides, useful for preventing,
 diagnosing or treating NOVA-associated disorders e.g. diabetes, cancer,
 Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
 disorders, and asthma.

Claim 16; Page 138-142; 619pp; English.

The present invention relates to new isolated proteins (NOVA) and their
 coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
 any number from 1 to 48. The NOVA proteins and coding sequences are
 useful in the manufacture of a medicament for treating a syndrome
 associated with a human disease, preferably a NOVA-associated disorder.
 The NOVA coding sequences and proteins are useful for treating,
 preventing or diagnosing diseases such as metabolic disorders, diabetes,
 obesity, infectious disease, anorexia, cancer-associated cachexia,
 cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
 disease, immune disorders, haematopoietic disorders, cardiovascular
 disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
 disturbances associated with obesity, metabolic syndrome X or wasting
 disorders associated with chronic diseases or various cancers. The NOVA
 coding sequences and proteins may also be used as targets for the
 identification of small molecules that modulate or inhibit e.g.
 neurogenesis, cell differentiation, cell proliferation, haematopoiesis,
 wound healing and angiogenesis, in gene therapy, in generation of
 antibodies that bind immunospecifically to NOVA substances for use in
 therapeutic or diagnostic methods

Sequence 14061 BP; 2854 A; 4029 C; 4839 G; 2339 T; 0 U; 0 Other;

Query Match 100.0%; Score 401; DB 6; Length 14061;
 Best Local Similarity 100.0%; Pred. No. 2e-86;

XX Human; ss; gene; serine/threonine kinase; protein kinase; 12599;
 KW cardiovascular disease; heart failure; myocardial infarction;
 KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma;
 KW blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
 KW haemolytic anaemia; cellular proliferative disorder; cancer;
 KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
 KW psoriasis; inflammatory bowel disease; rheumatoid arthritis;
 KW multiple sclerosis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..71
 FT /*tag= a
 FT CDS 72..23978
 FT /*tag= b
 FT /*product= "kinase 12599"
 FT /*note= "This CDS is specifically claimed in claim 2"
 FT 3'UTR 23979..24120
 FT /*tag= c
 XX
 US2002168742-A1.
 XX
 PD 14-NOV-2002.
 XX
 PF 15-FEB-2002; 2002US-00077130.
 XX
 PR 15-FEB-2001; 2001US-0269201P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Kapeller-Libermann R, Acton SL;
 XX
 DR WPI; 2003-298729/29.
 DR P-PSDB; ABG76187.
 XX
 PT Novel isolated human protein kinase, designated 59079 or 12599
 PT polypeptide, useful as diagnostic and therapeutic agents for preventing
 PT cardiovascular diseases, proliferative disorders, and protein kinase
 PT disorders.
 XX
 PS Claim 2; Page 58-84; 119pp; English.
 XX
 CC The invention relates to an isolated human serine/threonine or protein
 CC kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule
 CC comprising at least 85% identity to the nucleic acids appearing as
 CC ABX11641 and ABX11642 or their complement, a naturally occurring variant
 CC of the kinases or their fragments. Also included are a non-human host
 CC cell containing the nucleic acids, an antibody specific for the proteins,
 CC identifying a compound which binds to the kinase (by contacting the
 CC kinase or a cell expressing the kinase with a test compound and
 CC determining whether the kinase binds to the test compound) and modulating
 CC the activity of kinase using the identified compound. The kinases and
 CC their encoding nucleic acids are useful as diagnostic and therapeutic
 CC agents for preventing a disease or condition associated with an aberrant
 CC or unwanted 59079 or 12599 activity in a subject, including
 CC cardiovascular diseases such as heart failure, and myocardial infarction;
 CC disorders involving blood vessels such as atherosclerosis, and Kaposi's
 CC sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia,
 CC Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders
 CC such as cancer; and protein kinase disorders such as autoimmune
 CC disorders, diabetes mellitus, psoriasis, inflammatory bowel disease,
 CC rheumatoid arthritis, and multiple sclerosis (many examples of diseases
 CC and disorders are included in the specification). The kinases, their
 CC encoding nucleic acids and antibodies are useful in screening assays,
 CC detection assays (e.g. forensic biology), and predictive medicine (e.g.
 CC diagnostic assays, prognostic assays, and monitoring clinical trials and
 CC pharmacogenomics). The kinases and their encoding nucleic acids are
 CC useful as query sequences to perform a search against public databases to
 CC identify other family members or related sequences. The present sequence
 CC encodes the kinase 12599

SQ Sequence 24120 BP; 4719 A; 7265 C; 8278 G; 3858 T; 0 U; 0 Other;
 Query Match 100.0%; Score 401; DB 8; Length 24120;
 Best Local Similarity 100.0%; Pred. No. 2.2e-86;
 Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GAAGACCAAGTGGTGGCGGCCAGAGGAGACGTGAGCTGCGCTGTGAGCTGTCAAGGCG 60
 Db 8571 GAAGACCAAGTGGTGGCGGCCAGAGGAGGACGTGAGCTGCGCTGTGAGCTGTCAAGGCG 8630
 Qy 61 GGAAGCGCCGTGCACTGGCTGAAGGACAGGAGGAGCCATCCGCAAGAGCCAGAGTATGAT 120
 Db 8631 GGAAGCGCCGTGCACTGGCTGAAGGACAGGAGGAGCCATCCGCAAGAGCCAGAGTATGAT 8690
 Qy 121 GTGCTCTGCGAGGCGACGATGGCCATGCTGCTCATCCGCGGGGCTCGCTCAAGGACGCG 180
 Db 8691 GTGCTCTGCGAGGCGACGATGGCCATGCTGCTCATCCGCGGGGCTCGCTCAAGGACGCG 8750
 Qy 181 GGCAGGTACAGTGTGAGGTGGAGGCTTCCAAGAGCAAGCCAGCCCTCCATGTGGAAGAA 240
 Db 8751 GGCAGGTACAGTGTGAGGTGGAGGCTTCCAAGAGCAAGCCAGCCCTCCATGTGGAAGAA 8810
 Qy 241 AAAGCAAACTGCTTCACAGAGGAGCTGACCAATCTGAGGTGGAGGAGGAGGACACAGCT 300
 Db 8811 AAAGCAAACTGCTTCACAGAGGAGCTGACCAATCTGAGGTGGAGGAGGAGGACACAGCT 8870
 Qy 301 GTGTTACAGTGCAGAGCGAGCACCCCGCGGCCACAGTGACCTCGCGCAAGGGGCTCTTTG 360
 Db 8871 GTGTTACAGTGCAGAGCGAGCACCCCGCGGCCACAGTGACCTCGCGCAAGGGGCTCTTTG 8930
 Qy 361 GAGCTACGGGCTCAGGAGGAGCCAGCCAGCCAGCCAGGAGGG 401
 Db 8931 GAGCTACGGGCTCAGGAGGAGCCAGCCAGCCAGCCAGGAGGG 8971
 RESULT 7
 ABN21414
 ID ABN21414 standard; cDNA; 1005 BP.
 XX AC ABN21414;
 XX
 DT 24-JUN-2002 (first entry)
 DE Human ORFX polynucleotide sequence SEQ ID NO:11305.
 XX Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FN WO200192523-A2.
 XX
 PD 06-DEC-2001.
 XX
 PF 29-MAY-2001; 2001WO-US010836.
 XX
 PR 30-MAY-2000; 2000US-0206132P.
 XX
 PR 29-AUG-2000; 2000US-0228716P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shimkets RA, Leach MD;
 XX
 DR WPI; 2002-106308/14.
 DR P-PSDB; ABP05662.
 XX
 PT Novel human polypeptides and polynucleotides useful for diagnosing,

PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
PA (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Auendi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
DR P-PSDB; AAM41881.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
PT as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 5026; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: Immune system suppression,
CC Activ/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemias and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 3999 BP; 726 A; 1332 C; 1238 G; 703 T; 0 U; 0 Other;
Query Match 15.1%; Score 60.6; DB 4; Length 3999;
Best Local Similarity 54.8%; Pred. No. 9.5e-05;
Matches 120; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 10 TGGGTGGCGCCGAGGAGGACGTGGAGCTGCCTGTGAGCTGTACGGCGGGGACGCCC 69
Db 2612 TGTGTGGCCCTGGGGAGCCAGTGTGTGCTGAGCTGTGACTGTCCGGGCTGGCGCCCC 2553
QY 70 GTGCACCTGGCTGAAGGACGAGGAGCCATCCGCAAGAGCCAGAAATGATGTGTCTGC 129
Db 2552 GTGGTCTGGAGCCACAAATGGGAGGCCCGTGCAGGAGGGCGAGGGCTTAGAGCTCCATGCC 2493
QY 130 GAGGCGACGATGGCCATCTGTGTATCCCGGGGGCTCTCTCAAGAGCGCGGGCGAGTAC 189
Db 2492 GAGGCGCCCCCGCGAGTCTCTGTGCATCCAGGCTGCAGGCCCCAGCCCATGCAGGGCTCTAC 2433
QY 190 ACGTGTGAGGTGGAGGCTTCCAAAGAGCACAGGCAGCCTC 228
Db 2432 ACCTGCCAGTCTGGAGAGCCCGCCCGAGGCCCAAGCCTC 2394
RESULT 11
ADR67197
ID ADR67197 standard; DNA; 5378 BP.
XX
AC ADR67197;
XX
XX 02-DEC-2004 (first entry)
XX
XX Human bladder cancer associated nucleotide sequence.
XX
XX bladder cancer tissue; bladder cancer; cytostatic; gene; ds.
XX
XX Homo sapiens.
OS
XX WO2004076613-A2.
PN
XX

PD 10-SEP-2004.
XX
XX 24-FEB-2004; 2004WO-DE000364.
XX
XX 26-FEB-2003; 2003DE-01009729.
XX
XX (HERR/) HERR A.
PA (HINZ/) HINZMANN B.
PA (DAHL/) DAHL E.
PA (STAU/) STAUB E.
PA (PIIA/) PILARSKY C.
PA (SPEC/) SPECHT T.
XX
PI Herr A, Hinzmann B, Dahl E, Staub E, Pilarsky C, Specht T;
XX
XX WPI; 2004-653385/63.
XX
XX New nucleic acids, and encoded proteins, from bladder cancer tissue,
PT useful for diagnosis, treatment and in screening for specific binding
PT agents.
XX
XX Claim 1; Fig 3; 112pp; German.
XX
XX The present invention describes nucleic acids (I) associated with bladder
CC cancer tissue. Also described: (1) peptides and proteins (II) containing
CC an amino acid sequence encoded by (I); (2) a method for diagnosing
CC bladder cancer (BC), or monitoring its progression, that uses (I), (II)
CC or agents (Z) that inhibit, or bind to, (I) or (II), labelled with a
CC reporter; and (3) a method for treating BC that uses (I), (II) or (Z).
CC (I) and (II) have cytostatic activity. (I) and (II) can be used to detect
CC (and monitor progression of) bladder cancer (BC), or the risk of
CC developing it; to screen for specific binding agents (Z), and to treat
CC BC. (Z) are also useful as diagnostic and therapeutic agents. The present
CC sequence represents a human nucleotide sequence associated with bladder
CC cancer, which is used in the exemplification of the present invention.
XX
SQ Sequence 5378 BP; 940 A; 1696 C; 1799 G; 943 T; 0 U; 0 Other;
Query Match 15.1%; Score 60.6; DB 13; Length 5378;
Best Local Similarity 54.8%; Pred. No. 0.0001;
Matches 120; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 10 TGGGTGGCGCCGAGGAGGACGTGGAGCTGCCTGTGAGCTGTACGGCGGGGACGCCC 69
Db 2782 TGTGTGGCCCTGGGGAGCCAGTGTGTGCTGAGCTGTGAACCTGTCGGGCTGGCGCCCC 2841
QY 70 GTGCACCTGGCTGAAGGACGAGGAGCCATCCGCAAGAGCCAGAAATGATGTGTCTGC 129
Db 2842 GTGTCTGGAGCCACAAATGGGAGGCCCGTGCAGGAGGGCGAGGGCTTAGAGCTCCATGCC 2901
QY 130 GAGGCGACGATGGCCATCTGTGTATCCCGGGGGCTCTCGCTCAAGAGCGCGGGCGAGTAC 189
Db 2902 GAGGCGCCCCCGCGAGTCTCTGTGCATCCAGGCTGCAGGCCCCAGCCCATGCAGGGCTCTAC 2961
QY 190 ACGTGTGAGGTGGAGGCTTCCAAAGAGCACAGGCAGCCTC 228
Db 2962 ACCTGCCAGTCTGGAGAGCCCGCCCGAGGCCCAAGCCTC 3000
RESULT 12
ADD14722
ID ADD14722 standard; cDNA; 5382 BP.
XX
XX AC ADD14722;
XX
XX 01-JAN-2004 (first entry)
XX
XX Human src biomarker polynucleotide SEQ ID NO:116.
XX
XX predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
XX

OS Homo sapiens.
XX WO2003062395-A2.
XX PD 31-JUL-2003.
XX PF 17-JAN-2003; 2003WO-US001981.
XX PR 18-JAN-2002; 2002US-0350061P.
XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX PI Huang F, Fairchild CR, Lee FY, Shaw P;
XX DR WPI; 2003-636735/60.
XX DR P-PSDB; ADD14123.
XX PT New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX PS Claim 2; SEQ ID NO 116; 139pp; English.
XX CC The present invention describes a predictor set comprising a plurality of
CC polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.
XX SQ Sequence 5382 BP; 943 A; 1699 C; 1798 G; 942 T; 0 U; 0 Other;

Query Match 15.1%; Score 60.6; DB 10; Length 5382;
Best Local Similarity 54.8%; Pred. No. 0.0001;
Matches 120; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 10 TGGGTGGCGCCAGGGAGGACGTGGAGCTGCGCTGTGAGCTGTCCAGCGGGGGAACGCCCC 69
Db 2783 TGTGTGGCCCTGGGGAGCCAGTGGTGTGAGCTGTGAACTGTCCGGGCTGGCGCCCC 2842

QY 70 GTGCACCTGGCTGAGGACAGAGGCCATCCGCAAGAGCCGAGATGATGTGGTCTGC 129
Db 2843 GTGTCTGGAGCCCAATGGGAGGCCCTGTGAGGAGGCGGAGGCGCTTAGAGCTCCATGCC 2902

QY 130 GAGGGCACGATGGCCATGCTGGTCTATCCGCGGGGCCCTCGCTCAAGGACGCGGGCGAGTAC 189
Db 2903 GAGGGCCCCCGCCAGCTCTCTGTGATCAGGCTGACGGCCAGCCCATGAGGCTCTAC 2962

QY 190 ACGTGTGAGGTGGAGGCTTCCAGAGCACACGACCCCTC 228
Db 2963 ACCTGCCAGTCTGAGAGAGCCCCCGGAGCCCAAGCCCTC 3001

RESULT 13
AAK89083
ID AAK89083 standard; DNA; 20565 BP.
XX AC AAK89083;
XX DT 05-NOV-2001 (first entry)
XX XX Human digestive system antigen genomic sequence SEQ ID NO: 2659.
DE Human; digestive system antigen; gene therapy; cancer; appendicitis;
KW ulcerative colitis; infection; Hirschsprung's disease; chronic colitis;
KW digestive system disorder; Meckel's diverticulum; ds.
XX OS Homo sapiens.
XX PN WO200155314-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001324.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-MAR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214986P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.


```
PR 08-SEP-2000; 2000US-0233081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239933P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249224P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251858P.
PR 08-DEC-2000; 2000US-0251859P.
PR 08-DEC-2000; 2000US-0251860P.
PR 08-DEC-2000; 2000US-0251900P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Barash SC, Ruben SM;
XX WPI, 2001-502630/55.
XX
XX Polynucleotides encoding digestive system antigens, useful for
XX diagnosing, treating, preventing and/or prognosing disorders of the
XX digestive system, particularly cancer and cancer metastases.
XX
XX Disclosure; SEQ ID NO 2659; 986bp; English.
XX
XX The present invention provides the protein and coding sequences of a
XX number of human digestive system antigens. These can be used in the
XX diagnosis, treatment and prevention of digestive system disorders,
XX including cancer, Meckel's diverticulum, bacterial or parasitic
XX infections, appendicitis, Hirschsprung's disease, chronic colitis or
XX ulcerative colitis. The present sequence is a genomic DNA fragment
XX encoding a digestive system antigen of the invention
XX
XX Sequence 20565 BP; 4236 A; 6155 C; 5760 G; 4414 T; 0 U; 0 Other;
XX
XX Query Match 15.1%; Score 60.6; DB 4; Length 20565;
XX Best Local Similarity 54.8%; Pred. No. 0.00013;
XX Matches 120; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
XX
XX QY 10 TGGGTGGCGCCAGGGGAGGACGTGGAGCTGCGCTGTGAGCTGTGTCACGGCGGGACGCCC 69
XX Db 13230 TGTGTGGCCCTGGGAGCCAGTGTGCTGAGCTGTGAACCTGTCGGGGCTGGCGCCCC 13289
XX
XX QY 70 GTGCACTGGCTGAAGGACAGAGGCCATCCGAGAGCCAGAGATGATGTGTGCTGTC 129
XX Db 13290 GTGGTCTGGAGCCACAATGGAGGCCCGGTGAGAGGGCGAGGCTAGAGCTCCATGCC 13349
XX
XX QY 130 GAGGGCAGATGGCCATGCTGCTCATCCGGGGGCTCGCTCAAGGACGCGGGGAGTAC 189
XX Db 13350 GAGGGCCCCCGCCGAGTCTCTGCTCAGGCTGAGGCCAGGCCATGACAGGGGCTCTAC 13409
XX
XX QY 190 ACGTGTGAGTGGAGGCTTCCAGAGCAGCAGCCAGCCAGCCCTC 228
XX Db 13410 ACCTGCCAGTCTGGAGGAGCCCGCCGAGGCCCAAGCCCTC 13448
XX
XX RESULT 14
XX AAS28908/c
XX ID AAS28908 standard; DNA; 20565 BP.
XX
XX AC AAS28908;
XX
XX DT 07-NOV-2001 (first entry)
XX
XX Human immunoglobulin encoding genomic DNA SEQ ID No 270.
XX
XX Immunoglobulin; signal transduction pathway protein; cancer; ds;
XX antisense therapy; gene therapy; neurological disorder; renal disorder;
XX cardiovascular disorder; gastrointestinal disorder; pulmonary disorder;
XX reproductive disorder; immune system disorder; proliferative disorder;
XX muscular disorder.
XX
```


OS Homo sapiens.
XX WO200155315-A2.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001336.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229533P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236357P.
PR 29-SEP-2000; 2000US-0236388P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 06-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA

PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240360P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-465570/50.

Isolated nucleic acid molecule encoding a reproductive system antigen is
used in preventing, treating or ameliorating a medical condition.

Disclosure; SEQ ID NO 6045; 1297pp + Sequence Listing; English.

XX

The present invention provides the protein and coding sequences of a
number of human reproductive system related antigens. These can be used
in the prevention and treatment of reproductive system disorders,
including cancer. The present sequence is a genomic sequence encoding a
protein of the invention

SQ Sequence 20565 BP; 4236 A; 6155 C; 5760 G; 4414 T; 0 U; 0 Other;

Query Match

15.1%; Score 60.6; DB 4; Length 20565;

Best Local Similarity 54.8%; Pred. No. 0.00013;

Matches 120; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 10 TGGGTGGCGCCAGGGAGGACGTGGAGCTGGGCTGTGAGCTGTCCAGGGCGGGAACGCC 69

DB 13230 TGTGTGGCCCCCTGGGGAGCCAGTGTGCTGAGCTGTGAACCTGCCGGCTGGCGCCCC 13289

QY 70 GTGCACCTGGCTGAAAGGACAGGAAGCCATCCGCAAGAGCCAGAAAGTATGATGTGCTTGC 129

DB 13290 GTGGTCTGGAGCCACAATGGGAGCCCGTGCAGGAGGGCGAGGCGCTAGAGCTCCATGCC 13349

QY 130 GAGGGCAGGATGGCCATGCTGGTTCATCGCGGGGCGCTCGCTCAAGGAGCGGGCGAGTAC 189

DB 13350 GAGGGCCCCCGCCCGAGTCTTGCATCCAGGCTGCAGGCCCGCCATGCAGGGCTCTAC 13409

QY 190 ACGTGTGAGGTGGAGGCTTCCAGAGCACAGCCAGCCCTC 228

DB 13410 ACCTGCCAGTCTGGAGCAGCCCCCGGAGCCCCCAAGCCTC 13448

Search completed: March 20, 2005, 18:55:02

Job time : 234.766 secs

QY 241 CCATGGCGTTCTCCACCTGGACATPAAAGCCCTCTTAAATCTCTGATGTGTCATCTCTGCCG 300
DB 690 CCATGGCGTTCTCCACCTGGACATPAAAGCCCTCTTAAATCTCTGATGTGTCATCTGCCG 749
QY 301 GGAAGACATTAATAATCTGCACATTTGGCTTTGGCCAGAAATCACCCACAGAGCTGCA 360
DB 750 GGAAGA CATTAATAATCTGCACATTTGGCTTTGGCCAGAAATCACCCACAGAGCTGCA 809
QY 361 GTTCAGCCAGTACGGCTCCCTGAGTTCTGCTCCCGAGATCATCAGCAGAACCTGT 420
DB 810 GTTCAGCCAGTACGGCTCCCTGAGTTCTGCTCCCGAGATCATCAGCAGAACCTGT 869
QY 421 GAGGGAAGCTCCGACATTTGGGCCATGGGTGTCACTCTCTACCTCAGCTCAGCTGTCTC 480
DB 870 GAGCGAAGCTCCGACATTTGGGCCATGGGTGTCACTCTCTACCTCAGCTCAGCTGTCTC 929
QY 481 ATCCCCATTTGGCGGAGAGTGACCGTGGCCACCTCTGAAAGAGCGCCAAAGACTTCAACAAGC 600
DB 990 GTCATGGAGCAGCCCCATGGCTGCCCACTCAGCGAAGAGCGCCAAAGACTTCAACAAGC 1049
QY 601 TAGCTGACAGAGCCCCCTCAGGCCGGCTTAGTGGGCCAGTGCCTCTCCCAACCCCTG 660
DB 1050 TAGCTGACAGAGCCCCCTCAGGCCGGCTTAGTGGGCCAGTGCCTCTCCCAACCCCTG 1109
QY 661 GTTCTCTGAAATCCATCCCTGCGAGGAGGCCCACTTCAACCAAGAGAGCTCAAGTT 1169
DB 1110 GTTCTCTGAAATCCATCCCTGCGAGGAGGCCCACTTCAACCAAGAGAGCTCAAGTT 1169
QY 721 CTTCTGGCCGGAAGTCGTGCGAGCTTCCTCTGATGAGCTACAAAGTCCATCTCTGGTAT 780
DB 1170 CTTCTGGCCGGAAGTCGTGCGAGCTTCCTCTGATGAGCTACAAAGTCCATCTCTGGTAT 1229
QY 781 GCGCTCCATCCCTGAGCTGCTGGGGGGCCCAACGACAGCCCTCCCTCGGGGTAGCCCG 840
DB 1230 GCGCTCCATCCCTGAGCTGCTGGGGGGCCCAACGACAGCCCTCCCTCGGGGTAGCCCG 1289
QY 841 GCACCTCTGACGGACATGTGTGGCTCTCCAGTTTCCTCTCTCTCTCTGACACAGAGCT 900
DB 1290 GCACCTCTGACGGACATGTGTGGCTCTCCAGTTTCCTCTCTCTCTCTGACACAGAGCT 1349
QY 901 CGCCCCATTTGGCCGGCTAAGTCACTGCGACCTCTCCCGGTGACACACTCACCACTGCT 960
DB 1350 CGCCCCATTTGGCCGGCTAAGTCACTGCGACCTCTCCCGGTGACACACTCACCACTGCT 1409
QY 961 GCACCCCGGGCTTCTCTGGGCCCTCGGCCAGCTGCTCCGAGGGTGC CGGGCCACCGGCCG 1080
DB 1410 GCACCCCGGGCTTCTCTGGGCCCTCGGCCAGCTGCTCCGAGGGTGC CGGGCCACCGGCCG 1080
QY 1021 GCGCTCCACCGAGCCCGAGCTCCGCTGCTATCTCCGAGGGTGC CGGGCCACCGGCCG 1080
DB 1470 GCGCTCCACCGAGCCCGAGCTCCGCTGCTATCTCCGAGGGTGC CGGGCCACCGGCCG 1529
QY 1081 CCAGGGCTGCTGCTCCCGGACAGAGTCACTCCGACCTGTTTACCAACGAGCGGTGA 1140
DB 1530 CCAGGGCTGCTGCTCCCGGACAGAGTCACTCCGACCTGTTTACCAACGAGCGGTGA 1589
QY 1141 GAGCCCTGAGCAGCGGGCCCTGGCCCCCGGGAGCAGCGGCACCCGGCCCGGGGGGGCA 1200
DB 1590 GAGCCCTGAGCAGCGGGCCCTGGCCCCCGGGAGCAGCGGCACCCGGCCCGGGGGGGCA 1649
QY 1201 CCTGCTGAGGGCGGTATATTGGGGGGCTGCCAGGCTGCGGAGCCTGATGGA 1260
DB 1650 CCTGCTGAGGGCGGTATATTGGGGGGCTGCCAGGCTGCGGAGCCTGATGGA 1709
QY 1261 GCACCGGTGCTGAGGAGGAGGCCCGCAGGAGGAGCAGGCCACCTCTCTGCGCAAGC 1320
DB 1710 GCACCGGTGCTGAGGAGGAGGCCCGCAGGAGGAGCAGGCCACCTCTCTGCGCAAGC 1769

QY 1321 CCCCTCATTTGAGAGTGCCTCCGGCTGCTGCTCTGGCACCCACACTTTGGGCCCTTGGCCA 1380
DB 1770 CCCCTCATTTGAGAGTGCCTCCGGCTGCTGCTCTGGCACCCACACTTTGGGCCCTTGGCCA 1829
QY 1381 CAGCCACTCCCTTGGAAATGACTCTCTCCGAGCACCCCCGCCCTCTCTCGAGGGCTCGGG 1440
DB 1830 CAGCCACTCCCTTGGAAATGACTCTCTCCGAGCACCCCCGCCCTCTCTCGAGGGCTCGGG 1889
QY 1441 TGAGGCAAGGAGTGCCTTTCAGCCCTTCGCGGGGGCCCTTATCAGGGACATGCGGCA 1500
DB 1890 TGAGGCAAGGAGTGCCTTTCAGCCCTTCGCGGGGGCCCTTATCAGGGACATGCGGCA 1949
QY 1501 CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGTGGCCACCCAGGCACTGCTCAGCCAGA 1560
DB 1950 CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGTGGCCACCCAGGCACTGCTCAGCCAGA 2009
QY 1561 GAGGCCATTCGCCGACAGCCCTTGGGGGACAGCCCTTCTGCGCACCCCAAGCAGGG 1620
DB 2010 GAGGCCATTCGCCGACAGCCCTTGGGGGACAGCCCTTCTGCGCACCCCAAGCAGGG 2069
QY 1621 TTCTGCCCCCAGAGGGCTGACAGCCCGCAGCCAGCCCTTCTGCGCACCCCAAGCAGGG 1680
DB 2070 TTCTGCCCCCAGAGGGCTGACAGCCCGCAGCCAGCCCTTCTGCGCACCCCAAGCAGGG 2129
QY 1681 CTTCCCTCCAGGATCTTGCAAAGAGGCGCCCTTAGTACCTCAAGCCCTTCTTGGGACA 1740
DB 2130 CTTCCCTCCAGGATCTTGCAAAGAGGCGCCCTTAGTACCTCAAGCCCTTCTTGGGACA 2189
QY 1741 GCCCCAGGCAACCCCTGCCCCCTGCAAGCAAGCCCCCAATTGGACTTAAAGATGGGGCC 1800
DB 2190 GCCCCAGGCAACCCCTGCCCCCTGCAAGCAAGCCCCCAATTGGACTTAAAGATGGGGCC 2249
QY 1801 TGAGAGCATCTCTCTTCTGGGAGGCCAAACCCGGCCCTGAGTTCCCGAGGGTCCAGC 1860
DB 2250 TGAGAGCATCTCTCTTCTGGGAGGCCAAACCCGGCCCTGAGTTCCCGAGGGTCCAGC 2309
QY 1861 CTCCAGCGAGCTCTTCCCAAGTGAAGTCCCTCAGGGTGGGCTCTCTCCAGGTGGGCAC 1920
DB 2310 CTCCAGCGAGCTCTTCCCAAGTGAAGTCCCTCAGGGTGGGCTCTCTCCAGGTGGGCAC 2369
QY 1921 AGAGCTTGGCCCTCTCTGATGCGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 1980
DB 2370 AGAGCTTGGCCCTCTCTGATGCGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 2429
QY 1981 CTCCACACCCACCTTTCAGCGGCTCAGGAAACAGGGGACCATGCGCAAGTTCTCCCTGGG 2040
DB 2430 CTCCACACCCACCTTTCAGCGGCTCAGGAAACAGGGGACCATGCGCAAGTTCTCCCTGGG 2489
QY 2041 TGGTCGCGGGGCTACGAGGGCTGCTGGCTATGCGACCTTTGCTTTGGTGGAGATGC 2100
DB 2490 TGGTCGCGGGGCTACGAGGGCTGCTGGCTATGCGACCTTTGCTTTGGTGGAGATGC 2549
QY 2101 AGGGGCACTGCTGGGGCAGGGCCCATGTTGGGCGAGGATAGCTTGGGCTGTGTCCAGTTC 2160
DB 2550 AGGGGCACTGCTGGGGCAGGGCCCATGTTGGGCGAGGATAGCTTGGGCTGTGTCCAGTTC 2609
QY 2161 GGAGGAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
DB 2610 GGAGGAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2669
QY 2221 CAGGGCTGAGAGCCCATGCTCCCGAGGCTGAGTCCAGTCCGAGGAGCAGCAGAGGAGG 2280
DB 2670 CAGGGCTGAGAGCCCATGCTCCCGAGGCTGAGTCCAGGAGGAGGAGGAGGAGGAGGAGG 2729
QY 2281 TCCCAACGAGGCTCTCCAGAGCCCAACCCATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340
DB 2730 TCCCAACGAGGAGGCTCTCCAGAGCCCAACCCATGGGAGGAGGAGGAGGAGGAGGAGGAGG 2789
QY 2341 GAGATTCGCGGACCTGTGAGGTGATGCGAGGCGGCCGACCAATATCCCTGGACATTTTC 2400
DB 2790 GAGATTCGCGGACCTGTGAGGTGATGCGAGGCGGCCGACCAATATCCCTGGACATTTTC 2849
QY 2401 CGAGGTGGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTT 2460

Db 2850 CGAGTGAACCCCGCTACCTCAACTCTCAGACCTGTACGATATCAGTACTCCCAT 2909
Qy 2461 CGAGTTTATGATCTTTCAGAAAGTCCCCAAGTCCGCTCAGCAGAGCCGCCCTCCCCCAT 2520
Db 2910 CGAGTTTATGATCTTTCAGAAAGTCCCCAAGTCCGCTCAGCAGAGCCGCCCTCCCCCAT 2969
Qy 2521 GCCTCAGGAGGAGCTGGCCGAGTCCCGAGGCCACGCTGGCCCTCGCCAGGTTGAAGTGGG 2580
Db 2970 GCCTCAGGAGGAGCTGGCCGAGTCCCGAGGCCACGCTGGCCCTCGCCAGGTTGAAGTGGG 3029
Qy 2581 CCCCCACGAGGAGCTGGAGATCACAGAGGAGTCAGAGGATGGAGCGCTCTCGGCAGA 2640
Db 3030 CCCCCACGAGGAGCTGGAGATCACAGAGGAGTCAGAGGATGGAGCGCTCTCGGCAGA 3089
Qy 2641 GCCTCGCTGGCAGAGGAGCGCAAGTGTCTTCGCGCTCACAGCGCTCTTCCACTTCCC 2700
Db 3090 GCCTCGCTGGCAGAGGAGCGCAAGTGTCTTCGCGCTCACAGCGCTCTTCCACTTCCC 3149
Qy 2701 TGGAGGAGCAGCTCGCGCTGGATGAGCGCTGCAGAGCTGGGGCTGCGTGAGAGTGAAGGC 2760
Db 3150 TGGAGGAGCAGCTCGCGCTGGATGAGCGCTGCAGAGCTGGGGCTGCGTGAGAGTGAAGGC 3209
Qy 2761 CTCCTGTGAGACATCTCCCGATCTCTGAAGGCCAGCGGAAAGTCTGGAGAAAGAGGG 2820
Db 3210 CTCCTGTGAGACATCTCCCGATCTCTGAAGGCCAGCGGAAAGTCTGGAGAAAGAGGG 3269
Qy 2821 GCCCCCCAGGAGAACCCAGGCGCTTCTTCTTCCGGCTCTCAGGCTCTGAAGAGCTGGGA 2880
Db 3270 GCCCCCCAGGAGAACCCAGGCGCTTCTTCTTCCGGCTCTCAGGCTCTGAAGAGCTGGGA 3329
Qy 2881 CCGAGCGCGCATCTTCTAAGGAGCTCTCAGATGAGAGTGTGTCTTGGGGCAGTCAAT 2940
Db 3330 CCGAGCGCGCATCTTCTAAGGAGCTCTCAGATGAGAGTGTGTCTTGGGGCAGTCAAT 3389
Qy 2941 GACACTGGCTGCCAGGTGTGAGCCAGCCAGCTGCCAGGCCACTGGAGCAAGAGCG 3000
Db 3390 GACACTGGCTGCCAGGTGTGAGCCAGCCAGCTGCCAGGCCACTGGAGCAAGAGCG 3449
Qy 3001 A 3001
Db 3450 A 3450

RESULT 2
AX039410
LOCUS AX039410 5007 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 3 from Patent WO0063381.
ACCESSION AX039410
VERSION AX039410.1 GI:11229478
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
Zeng, W., Stanton, L. and Kong, H.
AUTHORS
TITLE Mammalian protein with putative function in signal transduction
JOURNAL Patent: WO 0063381-A 3 26-OCT-2000;
SCIOS INC. (US)
FEATURES
location/Qualifiers
1..5007
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
97...4929
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC16625.1"
/db_xref="GI:11229479"
translation="MQVTIEDVQAOITGTAQFEATIEGDPOPSVTWYKDSVQLVDSTR
LSQQEGTYSIVLRHVASKDAGVYTCIAQNTGQVLCRAELLVLGDNEDRSEKQSH
RKLSFYEVEEIGRGVGFVKRVQHKGNKILCAKFKIPUNSRTRAQYRERDILAA

ORIGIN

Query Match 99.9%; Score 2999.4; DB 6; Length 5007;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 CCTACGAGCAGAACTCCGGCCAGGCATACAGGAGCGAGACATCTCTGGCCCGCTGAG 60
Db 522 CCTACGAGCAGAACTCCGGCCAGGCATACAGGAGCGAGACATCTCTGGCCCGCTGAG 581
Qy 61 CCACCCGCTGTACGGGGCTGTGGACCGCTTTGAGACCCGACAGACCTCATCTCAT 120
Db 582 CCACCCGCTGTGTACGGGGCTGTGGACCGCTTTGAGACCCGACAGACCTCATCTCAT 641
Qy 121 CTGGAGCTGTGTCTCATCCGAGGAGCTGTGTGACCGCTGTACAGAAAGGGCGTGTGAC 180
Db 642 CTGGAGCTGTGTCTCATCCGAGGAGCTGTGTGACCGCTGTACAGAAAGGGCGTGTGAC 701
Qy 181 GGAGCCGAGGTCAAGGTCTACATCCAGCAGCTGTGTGAGGGGTGTGACTACTCTGCACAG 240
Db 702 GGAGCCGAGGTCAAGGTCTACATCCAGCAGCTGTGTGAGGGGTGTGACTACTCTGCACAG 761
Qy 241 CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTTAACATCTCTGATGTGTGATCTCTGCCG 300
Db 762 CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTTAACATCTCTGATGTGTGATCTCTGCCG 821
Qy 301 GGAAGACATTAATAATCTCGACCTTTGGGCTTTGGCCAGAAACATCACCCAGCAGAGCTGCA 360
Db 822 GGAAGACATTAATAATCTCGACCTTTGGGCTTTGGCCAGAAACATCACCCAGCAGAGCTGCA 881
Qy 361 GTTCAGCAGTACGGCTCCCTCGATTCGTCTTCCCGCAGATTCATCCAGAGAACCTGT 420
Db 882 GTTCAGCAGTACGGCTCCCTCGATTCGTCTTCCCGCAGATTCATCCAGAGAACCTGT 941
Qy 421 GAGCGAAGCTCCGACATTTGGGCCATTTGGGCCATTTGGGTGTCTCTTCTTCTTCTTCTTCT 480
Db 942 GAGCGAAGCTCCGACATTTGGGCCATTTGGGTGTCTCTTCTTCTTCTTCTTCTTCTTCTTCT 1001
Qy 481 ATCCCCATTTGCCGGCAGAGTGAACGCTGCCACCTCTCTGAAAGCTCTTGGAGGGCGCT 540
Db 1002 ATCCCCATTTGCCGGCAGAGTGAACGCTGCCACCTCTCTGAAAGCTCTTGGAGGGCGCT 1061
Qy 541 GTTCAGCAGCAGCCCATGGCTGCCACCTCTCAGCGAAGACGCGCAAGACTTTCATCAAGGC 600
Db 1062 GTTCAGCAGCAGCCCATGGCTGCCACCTCTCAGCGAAGACGCGCAAGACTTTCATCAAGGC 1121
Qy 601 TAGCGTCAGAGAGAGCCCTCAGGCGCGCGCTAGTGTGCGGCCAGTGTGCTCTTCTCCACCCCTG 660

Db	1122	TACGCTGCAGAGGCCCTTCAGGCCGGCCTAGTCGGGCCCAGTGCCTCTCCCAACCCCTG	1181
QY	661	GTTCTCGAAATCCATGCTCTCGGAGAGGCCACCTTCATCAACACCAAGCAGCTCAAGTT	720
Db	1182	GTTCTCGAAATCCATGCTCTCGGAGAGGCCACCTTCATCAACACCAAGCAGCTCAAGTT	1241
QY	721	CCTCTCTGCCCGAAGTGCCTGGCAGCGTTCCTGATGAGCTACAAGTCCATCTCTGGTGAT	780
Db	1242	CCTCTCTGCCCGAAGTGCCTGGCAGCGTTCCTGATGAGCTACAAGTCCATCTCTGGTGAT	1301
QY	781	GCGCTCCATCCCTGAGCTGCTCGCGGGGCCACCCGACAGCCCTCCTCGCGGTAGCCCG	840
Db	1302	GCGCTCCATCCCTGAGCTGCTCGCGGGGCCACCCGACAGCCCTCCTCGCGGTAGCCCG	1361
QY	841	GCACCTCTGCAGGGACACTGGTGGTCTCCAGTTCTCTCTCTCTCTGACAAACGAGCT	900
Db	1362	GCACCTCTGCAGGGACACTGGTGGTCTCTCCAGTTCTCTCTCTCTCTGACAAACGAGCT	1421
QY	901	CGCCCAATTTCGCCGGGCTAAGTCACTGSCACCTCCCGCGGTGACACACTCACCACTGCT	960
Db	1422	CGCCCAATTTCGCCGGGCTAAGTCACTGSCACCTCCCGCGGTGACACACTCACCACTGCT	1481
QY	961	GCACCCCGGGGCTTCTGCGGCCCTCGGCCAGCTCCTTGAGAAACCGAGGCCAGTGA	1020
Db	1482	GCACCCCGGGGCTTCTGCGGCCCTCGGCCAGCTCCTTGAGAAACCGAGGCCAGTGA	1541
QY	1021	GCGCTCCACCGAGGCCCGCAGCTCCGCGCTGATCTCCGAGGGTCCGGGCCACCGGCCG	1080
Db	1542	GCGCTCCACCGAGGCCCGCAGCTCCGCGCTGATCTCCGAGGGTCCGGGCCACCGGCCG	1601
QY	1081	CCAGGGTGTGTCGCCCGGACAGCGTCACTCCGAGCCTGTTCTACACAGAGCGGTGA	1140
Db	1602	CCAGGGTGTGTCGCCCGGACAGCGTCACTCCGAGCCTGTTCTACACAGAGCGGTGA	1661
QY	1141	GAGCCCTGAGCAGGGGCCCTGGCCCGGGAGAGCGGCACCGGCCCGCGCGCGCA	1200
Db	1662	GAGCCCTGAGCAGGGGCCCTGGCCCGGGAGAGCGGCACCGGCCCGCGCGCGCA	1721
QY	1201	CCTGCTGAAGGGCGGGTACATTGCGGGGGCGCTGCCAGGCTCGCGGAGCACTGATGA	1260
Db	1722	CCTGCTGAAGGGCGGGTACATTGCGGGGGCGCTGCCAGGCTCGCGGAGCACTGATGA	1781
QY	1261	GCACCGGTCTGGAGAGAGCGCGCCAGGAGAGAGAGCCACCTCTGGCCAAAGC	1320
Db	1782	GCACCGGTCTGGAGAGAGCGCGCCAGGAGAGAGAGCCACCTCTGGCCAAAGC	1841
QY	1321	CCCTCATTCGAGACTGCCCTCCGGTGTCTGCAAGCCATGCGCCCTTGCGCA	1380
Db	1842	CCCTCATTCGAGACTGCCCTCCGGTGTCTGCAAGCCATGCGCCCTTGCGCA	1901
QY	1381	CAGGCATCCTCTGGAACTGCTCTCGAGACACCCCGCCCTCTCTCGAGGCTCTGGG	1440
Db	1902	CAGGCATCCTCTGGAACTGCTCTCGAGACACCCCGCCCTCTCTCGAGGCTCTGGG	1961
QY	1441	TGAGGCACAGCGACTGCTTACGCCCCCTCCGGGGGGCCCTATCAGGACATGGGCA	1500
Db	1962	TGAGGCACAGCGACTGCTTACGCCCCCTCCGGGGGGCCCTATCAGGACATGGGCA	2021
QY	1501	CCCTCAGGGTCCAAAGAGTTCATCCACTGCTGGGCCACCCAGGCACCTGCTCAGCCAGA	1560
Db	2022	CCCTCAGGGTCCAAAGAGTTCATCCACTGCTGGGCCACCCAGGCACCTGCTCAGCCAGA	2081
QY	1561	GAGGCCATCCCGGACAGCCCTTGGGGGAGCAGCCCTTCTGCAACCCCAACAGGG	1620
Db	2082	GAGGCCATCCCGGACAGCCCTTGGGGGAGCAGCCCTTCTGCAACCCCAACAGGG	2141
QY	1621	TTCTGCCCCCAGAGGGCTGAGCCGCCACCCAGCAGTTGGCCCATGCCCCCTCTGGGCTC	1680
Db	2142	TTCTGCCCCCAGAGGGCTGAGCCGCCACCCAGCAGTTGGCCCATGCCCCCTCTGGGCTC	2201
QY	1681	CTTCCCTCCAGATCTTGCAAGAGGCCCTTAGTACCTCAAGCCCTTCTTGGGACA	1740

[illegible]

QY	1861	CTCCACGGCAGCTCTTCCAAAGT	GAGTCCCTCAGGGTGGGCTCTCCACAGGTGGGCAC	1920	
Db	5303	CTCCAGGGCAGCTCTTCCAAAGT	GAGTCCCTCAGGGTGGGCTCTCCACAGGTGGGCAC	5362	
QY	1921	AGAGCCTGGCCCTCCCTCGATGCGGAGGGCTGGACCCAGAGGCTCAGGATCTGTCGGA	1980		
Db	5363	AGAGCCTGGCCCTCCCTCGATGCGGAGGGCTGGACCCAGAGGCTCAGGATCTGTCGGA	5422		
QY	1981	CTCCACA	CCCACTTTGCAGCGGCTTCAGGAACAGGCGCAACATGCGCAAGTTCTCCCTGGG	2040	
Db	5423	CTCCACA	CCCACTTTGCAGCGGCTTCAGGAACAGGCGCAACATGCGCAAGTTCTCCCTGGG	5482	
QY	2041	TGGTTCGGGGGGCTACGACGGGTGGCTGGCTATATGGCA	CTTTGGCTTTTGGTGGAGATGC	2100	
Db	5483	TGGTTCGGGGGGCTACGACGGGTGGCTGGCTATATGGCA	CTTTGGCTTTTGGTGGAGATGC	5542	
QY	2101	AGGGGGCATGCTGGGGCAGGGGCCATGTTGGGCCACGATAGCCTGGGCTGTGTCCCAGTC	2160		
Db	5543	AGGGGGCATGCTGGGGCAGGGGCCATGTTGGGCCACGATAGCCTGGGCTGTGTCCCAGTC	5602		
QY	2161	GAGAGAGAGGAGCAGAGGAGAGGCGCAGGGCTGATGCCAGT	TCGGAAGGAGCAGGAGGC	2220	
Db	5603	AGAGGAGAGGAGCAGAGGAGAGGCGCAGGGCTGATGCCAGT	TCGGAAGGAGCAGGAGGC	5662	
QY	2221	CAGGGCTGAGAGCCACATGCGCCCAAGTCAAGTCAAGGCTGTGGCTGAGGTGGCAGGGC	2280		
Db	5663	CAGGGCTGAGAGCCCACTGCCCCAGTCAAGTCAAGGCTGTGGCTGAGGTGGCAGGGC	5722		
QY	2281	TCCCA	CCAGGAGTCTCCAGAGCCACCCCATGGGAGGACATCGGGCAGGTCTCCCTGGT	2340	
Db	5723	TCCCA	CCAGGAGTCTCCAGAGCCACCCCATGGGAGGACATCGGGCAGGTCTCCCTGGT	5782	
QY	2341	GCAGATCCGGACCTGTTCAGTGTGCGGAGGGCGCGCACAAATATCCCTGGACAATTTC	2400		
Db	5783	GCAGATCCGGACCTGTTCAGTGTGCGGAGGGCGCGCACAAATATCCCTGGACAATTTC	5842		
QY	2401	CGAGTGGACCCCGCCTTACCTCAACCTCTCAGACCTGTGTACGATATCAAGTACTCCCATTT	2460		
Db	5843	CGAGTGGACCCCGCCTTACCTCAACCTCTCAGACCTGTGTACGATATCAAGTACTCCCATTT	5902		
QY	2461	CGAGTTTATGATCTTCAGGAAATCCGCCAAGTCCGCTCAGCCAGAGCGCGCTCCCGCAT	2520		
Db	5903	CGAGTTTATGATCTTCAGGAAATCCGCCAAGTCCGCTCAGCCAGAGCGCGCTCCCGCAT	5962		
QY	2521	GGCTGAGGAGGAGCTGGCCGAGTTCCCGGAGCCACGTGGCCCTTCGCCAGTGAATCTGGG	2580		
Db	5963	GGCTGAGGAGGAGCTGGCCGAGTTCCCGGAGCCACGTGGCCCTTCGCCAGTGAATCTGGG	6022		
QY	2581	CCCCACGCAAGGCTTGAGATCA	CAGAGGATGTGGA	CGCGCTGCTGGCAGA	2640
Db	6023	CCCCACGCAAGGCTTGAGATCA	CAGAGGATGTGGA	CGCGCTGCTGGCAGA	6082
QY	2641	GGCTGCCGTGGGACGGAAGCGCAAGTGGTCTTCGCCGTTCAGCGCTCTTCCCATTTCCC	2700		
Db	6083	GGCTGCCGTGGGACGGAAGCGCAAGTGGTCTTCGCCGTTCAGCGCTCTTCCCATTTCCC	6142		
QY	2701	TGGGAGGCA	CCTGCCCTGATGAGCTTCAGAGCTGGGGCTGCGTGGAGAGAGTGAAGGC	2760	
Db	6143	TGGGAGGCA	CCTGCCCTGATGAGCTTCAGAGCTGGGGCTGCGTGGAGAGTGAAGGC	6202	
QY	2761	CTCCGTGGAGCACATCTCCGGATCTTGAAAGGACGGCGGAAGTCTGGAGAGGAGGG	2820		
Db	6203	CTCCGTGGAGCACATCTCCGGATCTTGAAAGGACGGCGGAAGTCTGGAGAGGAGGG	6262		
QY	2821	GCOCOCAGGAAGAACCCAGGCTTGTCTTCTTCGGGTCTCAGGTCTGAAGAGCTGGGA	2880		
Db	6263	GCOCOCAGGAAGAACCCAGGCTTGTCTTCTTCGGGTCTCAGGTCTGAAGAGCTGGGA	6322		
QY	2881	CCGAGCGCCGACATTCCTTAAGGAGTCTCAGATGAGACTGTGTCTCGGCCAGTCAGT	2940		
Db	6323	CCGAGCGCCGACATTCCTTAAGGAGTCTCAGATGAGACTGTGTCTCGGCCAGTCAGT	6382		

QY 2941 GACACTGGCTGCCAGGTGTTCAGCCAGCCAGCTGCCAGGCCACCTGGAGCAAGACGG 3000
Db 6383 GACACTGGCTGCCAGGTGTTCAGCCAGCCAGCTGCCAGGCCACCTGGAGCAAGACGG 6442

QY 3001 A 3001
Db 6443 A 6443

RESULT 4
LOCUS AR255867
DEFINITION Sequence 1 from patent US 6482624.
ACCESSION AR255867
VERSION AR255867.1 GI:27305065
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 5207)
AUTHORS Wei, M.-H., Kechum, K.A., Di Francesco, V. and Beasley, B.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins, and uses thereof
JOURNAL Patent: US 6482624-A 1 19-NOV-2002;
FEATURES Location/Qualifiers
source 1..5207
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 99.9%; Score 2997.8; DB 6; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;
Matches 2999; Conservative 0;

QY 1 CCTACGAGCAGAACTCGGGCCAGGCATACAGGAGGAGAGACATCTCTGGCCGCGCTGAG 60
Db 699 CCTACGAGCAGAACTCGGGCCAGGCATACAGGAGGAGAGACATCTCTGGCCGCGCTGAG 758

QY 61 CCACCCGCTGTACGGGGCTGTGGACAGAGTTGAGACCCGCAAGACCTCATCTCAT 120
Db 759 CCACCCGCTGTGTACGGGGCTGTGGACAGAGTTGAGACCCGCAAGACCTCATCTCAT 818

QY 121 CCTGAGCTGTCTCATCCGAGGAGCTGTGGACCGCTGTACAGAGAGGGGCTGGTAC 180
Db 819 CCTGAGCTGTCTCATCCGAGGAGCTGTGGACCGCTGTACAGAGAGGGGCTGGTAC 878

QY 181 GGAGGCCAGAGTCAAGGTCTACATCCAGCAGCTGGTGGAGGGGCTGCACTACTTCACAG 240
Db 879 GGAGGCCAGAGTCAAGGTCTACATCCAGCAGCTGGTGGAGGGGCTGCACTACTTCACAG 938

QY 241 CCATGGCGTTTCCACCTGGAGCATAAAGCCCTTAAACATCTCTGATGGTGCACTCTGCCG 300
Db 939 CCATGGCGTTTCCACCTGGAGCATAAAGCCCTTAAACATCTCTGATGGTGCACTCTGCCG 998

QY 301 GGAAGACATTAATCTGGAGCTTTGGCTTTGCCGAGACATCAACCCAGCAGAGCTGCA 360
Db 999 GGAAGACATTAATCTGGAGCTTTGGCTTTGCCGAGACATCAACCCAGCAGAGCTGCA 1058

QY 361 GTTCAGCAGTACGGCTCCCTTGGTTCGTCTCTCCCGAGATCATCCAGCAGAACCTGT 420
Db 1059 GTTCAGCAGTACGGCTCCCTTGGTTCGTCTCTCCCGAGATCATCCAGCAGAACCTGT 1118

QY 421 GAGCGAAGCTCCGACATTTGGGCCATGGGTGTCTATCTCTACCTCAGCTGACCTGCTC 480
Db 1119 GAGCGAAGCTCCGACATTTGGGCCATGGGTGTCTATCTCTACCTCAGCTGACCTGCTC 1178

QY 481 ATCCCCATTGGCCGCGAGAGTGACGTGGCCACCTCTCTGAACGTCTGGAGGGGGGT 540
Db 1179 ATCCCCATTGGCCGCGAGAGTGACGTGGCCACCTCTCTGAACGTCTGGAGGGGGGT 1238

QY 541 GTTCATGGAGCAGCCCATGGCTGCCACCTCAGCGAAGACGCCAAGACTTTCATCAAGGC 600
Db 1239 GTTCATGGAGCAGCCCATGGCTGCCACCTCAGCGAAGACGCCAAGACTTTCATCAAGGC 1298

QY 601 TACGCTGCAGAGAGCCCTCTCAGCCCGGCTAGTGGGGCCAGTGCCTCTCCACCCCTG 660
Db 1299 TACGCTGCAGAGAGCCCTCTCAGCCCGGCTAGTGGGGCCAGTGCCTCTCCACCCCTG 1358

QY 661 GTTCTGAAATCCATGCTCGGAGAGGAGCCACCTTTCATCAACCAAGAGCTCAAGTT 720
Db 1359 GTTCTGAAATCCATGCTCGGAGAGGAGCCACCTTTCATCAACCAAGAGCTCAAGTT 1418

QY 721 CTTCTGGCCGAAAGTCTGCTGGCAGCGTTCTCTGATGAGCTACAAAGTCAATCTGGTGAT 780
Db 1419 CTTCTGGCCGAAAGTCTGCTGGCAGCGTTCTCTGATGAGCTACAAAGTCAATCTGGTGAT 1478

QY 781 GCGCTTCATCTCCTGAGCTGCTGGGGCCACCGACAGACCCCTCCCTCGGCGTAGCCG 840
Db 1479 GCGCTTCATCTCCTGAGCTGCTGGGGCCACCGACAGACCCCTCCCTCGGCGTAGCCG 1538

QY 841 GCACCTCTGCAGGAGACACTGGTGGCTCTCTCAGTTCTCTCTCTCTGACCAAGAGCT 900
Db 1539 GCACCTCTGCAGGAGACACTGGTGGCTCTCTCAGTTCTCTCTCTCTGACCAAGAGCT 1598

QY 901 CGCCCATTTGCCCGGGCTAAGTCACTGCCACCTCTCCCGGTGACACACTCACACTGCT 960
Db 1599 CGCCCATTTGCCCGGGCTAAGTCACTGCCACCTCTCCCGGTGACACACTCACACTGCT 1658

QY 961 GCACCCCGGGGGCTTCTCGGGCCCTCGGCAGCTGCTGAGGAGCGGAGCCAGTGA 1020
Db 1659 GCACCCCGGGGGCTTCTCGGGCCCTCGGCAGCTGCTGAGGAGCGGAGCCAGTGA 1718

QY 1021 GCGCTCCACCGAGGCCCTCAGCTCGCCCTGCATCTCCGAGGGTGGCGGCCACCGGCCG 1080
Db 1719 GCGCTCCACCGAGGCCCTCAGCTCGCCCTGCATCTCCGAGGGTGGCGGCCACCGGCCG 1778

QY 1081 CAGGGCTGCTGCGCCCGGCAAGGTCACTCGCAGCTGTCTTACCAACAGGCGGGTGA 1140
Db 1779 CAGGGCTGCTGCGCCCGGCAAGGTCACTCGCAGCTGTCTTACCAACAGGCGGGTGA 1838

QY 1141 GAGCCCTGAGCAGCGGGCCCTGGCCCGGGGAGAGCGGCAACCGGCCCGGCGGCA 1200
Db 1839 GAGCCCTGAGCAGCGGGCCCTGGCCCGGGGAGAGCGGCAACCGGCCCGGCGGCA 1898

QY 1201 CTTGCTAAGGGCGGGTACATTGGCGGGGCGCTGCGAGGCTGCGCGAGCCACTGATGA 1260
Db 1899 CTTGCTAAGGGCGGGTACATTGGCGGGGCGCTGCGAGGCTGCGCGAGCCACTGATGA 1958

QY 1261 GCACCGCTGTGGAGGAGGCGCGCCAGGAGGAGCAGCCCTCTCTGGCCAAAGC 1320
Db 1959 GCACCGCTGTGGAGGAGGCGCGCGCCAGGAGGAGCAGCCCTCTCTGGCCAAAGC 2018

QY 1321 CCCCTCATTCAGACTGCGCTCGGCTGCTGCTGCGACCCCACTTGGGCCCTTGGCA 1380
Db 2019 CCCCTCATTCAGACTGCGCTCGGCTGCTGCTGCGACCCCACTTGGGCCCTTGGCA 2078

QY 1381 CAGCCACTCTCTGGAACATGACTCTCCGAGCAGACCCCGCCCTCTCTCGAGGCTCGCG 1440
Db 2079 CAGCCACTCTCTGGAACATGACTCTCCGAGCAGACCCCGCCCTCTCTCGAGGCTCGCG 2138

QY 1441 TGAGSCACAGCAGCTGCTTTCAGCCCTCTCGGGGGGGCCCTTATCAGGAGCATGGGCA 1500
Db 2139 TGAGSCACAGCAGCTGCTTTCAGCCCTCTCGGGGGGGCCCTTATCAGGAGCATGGGCA 2198

QY 1501 CCCTCAGGGCTCCAAAGCAGCTTCATCATCTGCTGGCCACCCAGGAGCTGCTCAGCAGA 1560
Db 2199 CCCTCAGGGCTCCAAAGCAGCTTCATCATCTGCTGGCCACCCAGGAGCTGCTCAGCAGA 2258

QY 1561 GAGGCATCTCCCGGAGACCCCTTGGGGGAGCGCCCTTCTGCGCACCCCAAGCAGGG 1620
Db 2259 GAGGCATCTCCCGGAGACCCCTTGGGGGAGCGCCCTTCTGCGCACCCCAAGCAGGG 2318

QY 1621 TTCTGCCCCCGCAGGAGGGCTGCGAGCCCGCCACCCAGCAGTTGCCCCATGCTCTGGCTC 1680
Db 2319 TTCTGCCCCCGCAGGAGGGCTGCGAGCCCGCCACCCAGCAGTTGCCCCATGCTCTGGCTC 2378

QY 1681 CTTCCCTCCAGGATCTTGCAAGAGGCCCCCTTATGATACCTCAAGCCCTTCTTGGACA 1740
DB 2379 CTTCCCTCCAGGATCTTGCAAGAGGCCCCCTTATGATACCTCAAGCCCTTCTTGGACA 2438
QY 1741 GCCCCAGGACCCCTGCCCCCTGCAAGAGAGGCCCCCTTATGAGTCTTAAAGTGGGCC 1800
DB 2439 GCCCCAGGACCCCTGCCCCCTGCAAGAGAGGCCCCCTTATGAGTCTTAAAGTGGGCC 2498
QY 1801 TGGAGACATCTCTCTCTCTGAGGAGCCCAAAACCCCGGCCCTGCACTTCCCAAGGTGACG 1860
DB 2499 TGGAGACATCTCTCTCTCTGAGGAGCCCAAAACCCCGGCCCTGCACTTCCCAAGGTGACG 2558
QY 1861 CTTCCAGGAGCTCTTCCAAAGTACGCTCCCTCAAGGTGGCTCTCTCAAGGTGGGAC 1920
DB 2559 CTTCCAGGAGCTCTTCCAAAGTACGCTCCCTCAAGGTGGCTCTCTCAAGGTGGGAC 2618
QY 1921 AGAGCTGGCCCCCTCTGATGCGAGGCTGACCCAGGAGGCTGAGGATCTGTCCGA 1980
DB 2619 AGAGCTGGCCCCCTCTGATGCGAGGCTGACCCAGGAGGCTGAGGATCTGTCCGA 2678
QY 1981 CTCCACACCCACTTTGCGAGGCTCAGGAAACAGGCGACCATGCGCAAGTTCTTCCCTGGG 2040
DB 2679 CTCCACACCCACTTTGCGAGGCTCAGGAAACAGGCGACCATGCGCAAGTTCTTCCCTGGG 2738
QY 2041 TGGTCGCGGGGCTACGAGGCTGGCTGCTATGCGACCTTTGCTTTGGTGAGATGC 2100
DB 2739 TGGTCGCGGGGCTACGAGGCTGGCTGCTATGCGACCTTTGCTTTGGTGAGATGC 2798
QY 2101 AGGGGGCATGTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGCCAGTC 2160
DB 2799 AGGGGGCATGTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGCCAGTC 2858
QY 2161 GGAGGAGGAGGAGGAGGAGGCGGCGAGGCTGAGTCCAGTGGAGGAGGAGGAGGCG 2220
DB 2859 GGAGGAGGAGGAGGAGGAGGCGGCGAGGCTGAGTCCAGTGGAGGAGGAGGAGGCG 2918
QY 2221 CAGGCTGAGAGCCACTGCCCCAGGTGAGTGCAGAGGCTGTGCTGAGGTGGGAGGGC 2280
DB 2919 CAGGCTGAGAGCCACTGCCCCAGGTGAGTGCAGAGGCTGTGCTGAGGTGGGAGGGC 2978
QY 2281 TCCACAGGAGCTCTCCAGAGCCACCCCATGGAGGACATCGGGCAGGCTCCCTGGT 2340
DB 2979 TCCACAGGAGCTCTCCAGAGCCACCCCATGGAGGACATCGGGCAGGCTCCCTGGT 3038
QY 2341 GCAGATCCGGACCTGTGAGGTGATGCGGAGGCGGCGACCAATATCCTTGGACATTC 2400
DB 3039 GCAGATCCGGACCTGTGAGGTGATGCGGAGGCGGCGACCAATATCCTTGGACATTC 3098
QY 2401 CGAGGTGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
DB 3099 CGAGGTGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 3158
QY 2461 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCGCCCTCCCCAT 2520
DB 3159 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCGCCCTCCCCAT 3218
QY 2521 GGCTGAGGAGGCTGGCCGATTCGCGAGCCCAAGTGGCCCTGGCCAGGTGAACCTGGG 2580
DB 3219 GGCTGAGGAGGCTGGCCGATTCGCGAGCCCAAGTGGCCCTGGCCAGGTGAACCTGGG 3278
QY 2581 CCCCCAGGAGGCTGGAGATCAAGAGGAGTCAAGAGATGAGAGGCTGGGAGGCTGGGAGA 2640
DB 3279 CCCCCAGGAGGCTGGAGATCAAGAGGAGTCAAGAGATGAGAGGATGGAAGGCTGGGAGA 3338
QY 2641 GGCTGCGGTGGGAGGAGGAGGCTGCTGCGCGTCAAGGAGGCTTTCACCTTCCC 2700
DB 3339 GGCTGCGGTGGGAGGAGGAGGCTGCTGCGCGTCAAGGAGGCTTTCACCTTCCC 3398
QY 2701 TGGAGGAGCCTGCGGCTGAGTGCAGAGCTGGGCTGGGCTGAGAGGAGTGAAGGC 2760
DB 3399 TGGAGGAGCCTGCGGCTGAGTGCAGAGCTGGGCTGGGCTGAGAGGAGTGAAGGC 3458
QY 2761 CTCCGTGGAGGACATCTCCCGATCTTGAAGGGCAGGCGGAGGCTGTGGAGAGGAGG 2820

DB 3459 CTCCGTGGAGCACATCTCCCGATCTCAAGGAGGCGCGAAGGTCTGGAGAGGAGG 3518
QY 2821 GCCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2880
DB 3519 GCCCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3578
QY 2881 CCGAGGCGGAGATCTTCTTAAGGAGGAGCTCTCAGATGAGACTGTGTGCTTGGGCCAGTCA 2940
DB 3579 CCGAGGCGGAGATCTTCTTAAGGAGGAGCTCTCAGATGAGACTGTGTGCTTGGGCCAGTCA 3638
QY 2941 GACACTGGCTGCGAGGTGTCAGGCCAGCCAGCTGCCCCAGGCCACCTTGGAGCAAGAGCG 3000
DB 3639 GACACTGGCTGCGAGGTGTCAGGCCAGCCAGCTGCCCCAGGCCACCTTGGAGCAAGAGCG 3698
QY 3001 A 3001
DB 3699 A 3699
RESULT 5
AR442826
LOCUS AR442826 5207 bp DNA linear PAT 20-FEB-2004
DEFINITION Sequence 1 from patent US 6670164.
ACCESSION AR442826
VERSION AR442826.1 GI:42670270
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 5207)
AUTHORS Wei, M.-H., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.
TITLE Isolated human kinase proteins
JOURNAL Patent: US 6670164-A 1 30-DEC-2003;
FEATURES Location/Qualifiers
source 1..5207
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Query Match 99.9%; Score 2997.8; DB 6; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCTACGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTTGGCCGCGCTGAG 60
DB 699 CCTACGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTTGGCCGCGCTGAG 758
QY 61 CCACCCGCTGGTCAAGGCTGTGGACCAAGTTGAGACCGCGAAGACCCCTCATCTCAT 120
DB 759 CCACCCGCTGGTCAAGGCTGTGGACCAAGTTGAGACCGCGAAGACCCCTCATCTCAT 818
QY 121 CTTGGAGCTGTGCTCATCCGAGGAGCTGTGGACCGGCTGTACAGGAGGCGGTGGTAC 180
DB 819 CTTGGAGCTGTGCTCATCCGAGGAGCTGTGGACCGGCTGTACAGGAGGCGGTGGTAC 878
QY 181 GGAGCGCGAGGTCAAGGTCTTACATCCAGCAGCTGGTGGAGGGGCTGCACTACCTGCA 240
DB 879 GGAGCGCGAGGTCAAGGTCTTACATCCAGCAGCTGGTGGAGGGGCTGCACTACCTGCA 938
QY 241 CCATGGGCTTCTCCACTGGACATTAAGCCCTCTAACATCTGTATGGTGCATCTCGCCG 300
DB 939 CCATGGGCTTCTCCACTGGACATTAAGCCCTCTAACATCTGTATGGTGCATCTCGCCG 998
QY 301 GGAAGACATTAATAATCTGGACATTTGGCTTTGGCCAGAACATCACCCACAGAGCTGCA 360
DB 999 GGAAGACATTAATAATCTGGACATTTGGCTTTGGCCAGAACATCACCCACAGAGCTGCA 1058
QY 361 GTTCAGCAGTACGGCTCCCTTGTGTTGCTTCCCGGAGATCATCCAGCAGAACCTGT 420
DB 1059 GTTCAGCAGTACGGCTCCCTTGTGTTGCTTCCCGGAGATCATCCAGCAGAACCTGT 1118
QY 421 GAGCGAAGCCCTCCGACATTTGGGCCATTTGGGGCTGTGCTCTCTACCTCAGCCTGACTC 480

Db 1119 GAGCGAAGCCTCCGACATTTGGGCCATTTGGGTGTGATCTCTACCTCAGCCTGACCTGCTC 1178
Qy 481 ATCCCAATTTGCGCGGAGAGTGACCGTGCACCTCTCTGAAACGTCTCTGGAGGGCGCGT 540
Db 1179 ATCCCAATTTGCGCGGAGAGTGACCGTGCACCTCTCTGAAACGTCTCTGGAGGGCGCGT 1238
Qy 541 GTCTGAGGACGCCCCCATTTGGTGCCTCAGCGAAGACGCCAAGACTTCATCAAGGC 600
Db 1239 GTCTGAGGACGCCCCCATTTGGTGCCTCAGCGAAGACGCCAAGACTTCATCAAGGC 1298
Qy 601 TACGCTGAGAGAGCCCTCAGGCGCGCGCTAGTGCAGCCCAAGTGCCTCTCCACCCCTG 660
Db 1299 TACGCTGAGAGAGCCCTCAGGCGCGCGCTAGTGCAGCCCAAGTGCCTCTCCACCCCTG 1358
Qy 661 GTTCCTGAAATCCATGCTGCGGAGAGGCCCATTTTCATCAACCAAGCAGCTCAAGTT 720
Db 1359 GTTCCTGAAATCCATGCTGCGGAGAGGCCCATTTTCATCAACCAAGCAGCTCAAGTT 1418
Qy 721 CCTCTGCGCGAAGTCTGCGAGCGTTCTCTGATGAGCTACAAGTCCATCTGTTGAT 780
Db 1419 CCTCTGCGCGAAGTCTGCGAGCGTTCTCTGATGAGCTACAAGTCCATCTGTTGAT 1478
Qy 781 CGCTCTCATCCTGAGCTGCTGCGGGGCCACCCGACAGCCCTCTCTCGCGTAGCCG 840
Db 1479 CGCTCTCATCCTGAGCTGCTGCGGGGCCACCCGACAGCCCTCTCTCGGGTAGCCG 1538
Qy 841 GCACCTCTGCAAGGACACTGTTGGCTCTCTCAGTTCTCTCTCTCTGCAACAGAGCT 900
Db 1539 GCACCTCTGCAAGGACACTGTTGGCTCTCTCAGTTCTCTCTCTCTCTGCAACAGAGCT 1598
Qy 901 CGCCCTTATGCGCGGGCTAAGTCTACTGCGACCTCTCCCGGTGACACACTCACCACTGCT 960
Db 1599 CGCCCTTATGCGCGGGCTAAGTCTACTGCGACCTCTCCCGGTGACACACTCACCACTGCT 1658
Qy 961 GCACCCCGGGCTTCTGCGGCTCTGCGGCTCTGAGGAGCGAGGCGCAGTGA 1020
Db 1659 GCACCCCGGGCTTCTGCGGCTCTGCGGCTCTGAGGAGCGAGGCGCAGTGA 1718
Qy 1021 CGCTCTCAACGAGGCCCTGCGCTGCACTCTCCGAGGGTGCGGGCCACCGCGCGC 1080
Db 1719 CGCTCTCAACGAGGCCCTGCGCTGCACTCTCCGAGGGTGCGGGCCACCGCGCGC 1778
Qy 1081 CCAGGGCTGCTGCGCGCGCAGCGTCTATCGCGAGCTGTTCTACCAAGCGGGGTGA 1140
Db 1779 CCAGGGCTGCTGCGCGCGCAGCGTCTATCGCGAGCTGTTCTACCAAGCGGGGTGA 1838
Qy 1141 GAGCCTTGAGCAGCGGGCCCTGGCCCGGAGCAGGCGCACCGCGCCCGCGCGCGCA 1200
Db 1839 GAGCCTTGAGCAGCGGGCCCTGGCCCGGAGCAGGCGCACCGCGCCCGCGCGCGCA 1898
Qy 1201 CTTGCTGAAGGGCGGGTACATTTGGGGGGCGCTGCGAGGCCCTGCGCGAGCCACTGATGGA 1260
Db 1899 CTTGCTGAAGGGCGGGTACATTTGGGGGGCGCTGCGAGGCCCTGCGCGAGCCACTGATGGA 1958
Qy 1261 GCACCGCGTCTGAGGAGGAGGCGCGCAGGAGGAGCAGGCGACCTCTCTGGCCAAAGC 1320
Db 1959 GCACCGCGTCTGAGGAGGAGGCGCGCAGGAGGAGCAGGCGACCTCTCTGGCCAAAGC 2018
Qy 1321 CCCCTCATTCGAGACTGCGCTCGGCTGCGCTCTGGCACCCACTTTGGCCCTGGCGCA 1380
Db 2019 CCCCTCATTCGAGACTGCGCTCGGCTGCGCTCTGGCACCCACTTTGGCCCTGGCGCA 2078
Qy 1381 CAGCGACTCTCTGGAACATGACTCTCTCGAGCACCCCGCGCTCTCTCGAGGGCTCGCG 1440
Db 2079 CAGCGACTCTCTGGAACATGACTCTCTCGAGCACCCCGCGCTCTCTCGAGGGCTCGCG 2138
Qy 1441 TGAGGCAAGGCACTGCTTCAAGCCCTCTCGGGGGGGCGCTATTCAGGAGCATGGGCA 1500
Db 2139 TGAGGCAAGGCACTGCTTCAAGCCCTCTCGGGGGGGCGCTATTCAGGAGCATGGGCA 2198
Qy 1501 CCCTCAGGGCTCCAGCAGCTTCCATCCACTGTTGGCCACCCAGGCACTGCTCAGCCAGA 1560

Db 2199 CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGA 2258
Qy 1561 GAGGCCATTCCTCCGACAGCCCTTGGGGGCAAGCCAGCCCTTTCTGCCACCCAAAGAGGG 1620
Db 2259 GAGGCCATTCCTCCGACAGCCCTTGGGGGCAAGCCAGCCCTTTCTGCCACCCAAAGAGGG 2318
Qy 1621 TTCTGCCCCCAGAGAGGGCTCAGCCGCCACCCAGCAGTTCGCCCATTCCTCTGGGCTC 1680
Db 2319 TTCTGCCCCCAGAGAGGGCTCAGCCGCCACCCAGCAGTTCGCCCATTCCTCTGGGCTC 2378
Qy 1681 CTTCCCTCCAGGACTTTGCAAGAGGCCCCCTTAGTACCTCTCAAGCCCTTTCTGGGACA 1740
Db 2379 CTTCCCTCCAGGACTTTGCAAGAGGCCCCCTTAGTACCTCTCAAGCCCTTTCTGGGACA 2438
Qy 1741 GCCCCAGGCAACCCCTCTGCTGCAAGAGGCCCCCCTTAGTACCTCTAAGAGTGGGGCC 1800
Db 2439 GCCCCAGGCAACCCCTCTGCTGCAAGAGGCCCCCCTTAGTACCTCTAAGAGTGGGGCC 2498
Qy 1801 TGGAGACATCTCTCTCTGGGAGGCCAAACCCGGGCCCTTGCAGTTCTCCCGAGGTCAGC 1860
Db 2499 TGGAGACATCTCTCTCTGGGAGGCCAAACCCGGGCCCTTGCAGTTCTCCCGAGGTCAGC 2558
Qy 1861 CTCCAGCGGAGCTCTTCCCAAGTGAGCTCCCTCAGGGTGGGCTCTCCAGGTGGGCAC 1920
Db 2559 CTCCAGCGGAGCTCTTCCCAAGTGAGCTCCCTCAGGGTGGGCTCTCCAGGTGGGCAC 2618
Qy 1921 AGAGCTTGGCCCCCTCTCTGATGCGGAGGCTGAGCCAGAGGCTAGGATCTGTTCGA 1980
Db 2619 AGAGCTTGGCCCCCTCTCTGATGCGGAGGCTGAGCCAGAGGCTAGGATCTGTTCGA 2678
Qy 1981 CTCCACACCCACCTTGCAGGGGCTCAGGACAGGCGACCATGCGCAAGTTCTCCCTGG 2040
Db 2679 CTCCACACCCACCTTGCAGGGGCTCAGGAAACAGGTGACCATGCGCAAGTTCTCCCTGG 2738
Qy 2041 TGGTTCGCGGGGCTACGCGAGCGTGGCTTGGCTATGCGACCTTTGCTTGGTGGAGATGC 2100
Db 2739 TGGTTCGCGGGGCTACGCGAGCGTGGCTTGGCTATGCGACCTTTGCTTGGTGGAGATGC 2798
Qy 2101 AGGGGGCATGCTGGGCGAGGGGCCATGTGGGCCAGGATAGCCTGGGCTGTGTCCAGTC 2160
Db 2799 AGGGGGCATGCTGGGCGAGGGGCCATGTGGGCCAGGATAGCCTGGGCTGTGTCCAGTC 2858
Qy 2161 GGAGGAGGAGCAGAGGAGGCGCAGGCTGAGTCCGAGTCCGAGGAGCAGCAGGAGGC 2220
Db 2859 GGAGGAGGAGCAGAGGAGGAGGCGCAGGCTGAGTCCGAGTCCGAGGAGCAGCAGGAGGC 2918
Qy 2221 CAGGGCTGAGAGGCCACTTGCCCCAGGTCAGTGAAGGCTGTGTGCTGAGGTGCGCAGGCG 2280
Db 2919 CAGGGCTGAGAGGCCACTTGCCCCAGGTCAGTGAAGGCTGTGTGCTGAGGTGCGCAGGCG 2978
Qy 2281 TCCACACGAGGCTCTCCAGAGCCCCACCCATGGGAGGACATCGGCGAGGCTCTCCCTGGT 2340
Db 2979 TCCACACGAGGCTCTCCAGAGCCCCACCCATGGGAGGACATCGGCGAGGCTCTCCCTGGT 3038
Qy 2341 GCAGATCGGGACCTGTGAGTGCAGGCGCGCGCGCGACACATATCCCTGGACATTTTC 2400
Db 3039 GCAGATCGGGACCTGTGAGTGCAGGCGCGCGCGCGCGACACATATCCCTGGACATTTTC 3098
Qy 2401 CGAGGTGAGCCCGCCCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
Db 3099 CGAGGTGAGCCCGCCCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 3158
Qy 2461 CGAGTTTATGATTTTCAAGGAAAGTCCCAGTCCGCTCAGCAGAGCGCGCCCTCCCCCAT 2520
Db 3159 CGAGTTTATGATTTTCAAGGAAAGTCCCAGTCCGCTCAGCAGAGCGCGCCCTCCCCCAT 3218
Qy 2521 GGCTGAGGAGGAGCTGCGCGAGTTCCGAGGCCACAGTGGCCCTTGGCCAGGTGAACCTGG 2580
Db 3219 GGCTGAGGAGGAGCTGCGCGAGTTCCGAGGCCACAGTGGCCCTTGGCCAGGTGAACCTGG 3278
Qy 2581 CCCCCAGCGGCTGAGATCACAGAGGAGTCAAGAGATGTGACGCGCTGTGGCAGA 2640
Db 3279 CCCCCAGCGGCTGAGATCACAGAGGAGTCAAGAGATGTGACGCGCTGTGGCAGA 3338


```
QY 1381 CAGCCACTCCCTGGAAACATGACTCTCCGAGACACCCCGCCCTCCTCGGAGGCTGCGG 1440
Db 2079 CAGCCACTCCCTGGAAACATGACTCTCCGAGACACCCCGCCCTCCTCGGAGGCTGCGG 2138
QY 1441 TGAGGACACAGGACTGCTTACGCCCCCTCGGGGGGGCCCCATATCAGGACATGGGGCA 1500
Db 2139 TGAGGACACAGGACTGCTTACGCCCCCTCGGGGGGGCCCCATATCAGGACATGGGGCA 2198
QY 1501 CCCTCAGGGCTCCAAAGACGCTTCCATCCATCTGGTGGCCACCCAGGACATGCTCAGCCAGA 1560
Db 2199 CCCTCAGGGCTCCAAAGACGCTTCCATCCATCTGGTGGCCACCCAGGACATGCTCAGCCAGA 2258
QY 1561 GAGGCAATCCCGGACAGCCCTTGGGGGACGAGCCCTTCTGCCACCCCAAGCAGGG 1620
Db 2259 GAGGCAATCCCGGACAGCCCTTGGGGGACGAGCCCTTCTGCCACCCCAAGCAGGG 2318
QY 1621 TTCTGCCCCCCAGGAGGCTGACGCCCCCACCAGCAGTTGCCCCATGCCCTCTCTGGCTC 1680
Db 2319 TTCTGCCCCCCAGGAGGCTGACGCCCCCACCAGCAGTTGCCCCATGCCCTCTCTGGCTC 2378
QY 1681 CTTCCCTCCAGATCTTCCAAAGAGGCCCCCTTAGTACCCTCAAGCCCCCTTCTTGGGACA 1740
Db 2379 CTTCCCTCCAGATCTTCCAAAGAGGCCCCCTTAGTACCCTCAAGCCCCCTTCTTGGGACA 2438
QY 1741 GCCCAGGACACCCCTGCCCCCTGCCAAAGACGCCCCCATTTGGACTCTAAGATGGGGCC 1800
Db 2439 GCCCAGGACACCCCTGCCCCCTGCCAAAGACGCCCCCATTTGGACTCTAAGATGGGGCC 2498
QY 1801 TGGAGACATCTCTTCTCTGGAGGCCCCAAACCCCGCCCCCTGCAGTTCCCCAGGGTCAGC 1860
Db 2499 TGGAGACATCTCTTCTCTGGAGGCCCCAAACCCCGCCCCCTGCAGTTCCCCAGGGTCAGC 2558
QY 1861 CTCACAGCGAGCTTCTCCAAAGTACGCTCCCTCAGGGTGGGCTCCTCCAGGTGGGCAC 1920
Db 2559 CTCACAGCGAGCTTCTCCAAAGTACGCTCCCTCAGGGTGGGCTCCTCCAGGTGGGCAC 2618
QY 1921 AGAGCTGGCCCCCTCCCTGGATGGGAGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 1980
Db 2619 AGAGCTGGCCCCCTCCCTGGATGGGAGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 2678
QY 1981 CTCACACCCACCTTGCAGCGGCTCAGGAAACAGGCGACCATGCGCAAGTTCTCCCTGGG 2040
Db 2679 CTCACACCCACCTTGCAGCGGCTCAGGAAACAGGCGACCATGCGCAAGTTCTCCCTGGG 2738
QY 2041 TGGTCGGGGGGCTACGAGGCGTGGCTGGCTATGGCACCTTTGGCTTTGGTGGAGATGC 2100
Db 2739 TGGTCGGGGGGCTACGAGGCGTGGCTGGCTATGGCACCTTTGGCTTTGGTGGAGATGC 2798
QY 2101 AGGGGGCATGTGGGGCAGGGGCCCATGCTGGGCCAGGATAGCCTGGGCTGTCTCCAGTC 2160
Db 2799 AGGGGGCATGTGGGGCAGGGGCCCATGCTGGGCCAGGATAGCCTGGGCTGTCTCCAGTC 2858
QY 2161 GGAGGAGGAGGACGAGGAGGGCCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2220
Db 2859 GGAGGAGGAGGACGAGGAGGGCCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2918
QY 2221 CAGGCTGAGAGCCCATCTGCCCCAGGTCAAGTCAAGAGGCTGTGCTGAGGTGCGGAGGGC 2280
Db 2919 CAGGCTGAGAGCCCATCTGCCCCAGGTCAAGTCAAGAGGCTGTGCTGAGGTGCGGAGGGC 2978
QY 2281 TCCACACAGGAGCTCTCAGAGGCCACCCCATGGAGGACATCGGGCAGGTCTCCCTGGT 2340
Db 2979 TCCACACAGGAGCTCTCAGAGGCCACCCCATGGAGGACATCGGGCAGGTCTCCCTGGT 3038
QY 2341 GCAGATCCGGACCTGTGAGGTGATGCGAGGCGGCCGACACATATCCCTGGACATTTTC 2400
Db 3039 GCAGATCCGGACCTGTGAGGTGATGCGAGGCGGCCGACACATATCCCTGGACATTTTC 3098
QY 2401 CGAGGTGAGCCCGGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
Db 3099 CGAGGTGAGCCCGGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 3158
QY 2461 CGAGTTTATGATCTTCAGGAAAGTCCCCCAAGTCCGCTCAGCCAGAGCGCCCTCCCCCAT 2520
```

```
Db 3159 CGAGTTTATGATCTTCAGGAAAGTCCCCAAGTCCGCTCAGCCAGAGCGCCCTCCTCCCAT 3218
QY 2521 GGCTGAGAGGAGCTGGCCGAGTTCCCGAGCCCACTGGGCCCTTGGCCAGAGTGAATCGG 2580
Db 3219 GGCTGAGAGGAGCTGGCCGAGTTCCCGAGCCCACTGGGCCCTTGGCCAGAGTGAATCGG 3278
QY 2581 CCCCCACGAGCCCTGGAGATCACAGAGGAGTCAAGAGATGTGGACGCGCTGCTGGCAGA 2640
Db 3279 CCCCCACGAGCCCTGGAGATCACAGAGGAGTCAAGAGATGTGGACGCGCTGCTGGCAGA 3338
QY 2641 GGCTGGCCCTGGGAGAGGAGCGCAAGTGGTCTCGCGCTCAGCAGCCTCTTCCACTTCCC 2700
Db 3339 GGCTGGCCCTGGGAGAGGAGCGCAAGTGGTCTCGCGCTCAGCAGCCTCTTCCACTTCCC 3398
QY 2701 TGGAGGACACTGCGCTGGATGAGCCCTGACAGCTGGGGCTGCTGAGAGAGTGAAGGC 2760
Db 3399 TGGAGGACACTGCGCTGGATGAGCCCTGACAGCTGGGGCTGCTGAGAGAGTGAAGGC 3458
QY 2761 CTCCGTGGAGCACATCTCCCGGATCTTGAAGGGCAGGCCGGAAGTCTGGAGAGGAGGG 2820
Db 3459 CTCCGTGGAGCACATCTCCCGGATCTTGAAGGGCAGGCCGGAAGTCTGGAGAGGAGGG 3518
QY 2821 GCCCCACAGGAAGAGCCAGGCTTGGTCTTCCCTCGGGCTCTCAGGTCTGAAGAGCTGGGA 2880
Db 3519 GCCCCACAGGAAGAGCCAGGCTTGGTCTTCCCTCGGGCTCTCAGGTCTGAAGAGCTGGGA 3578
QY 2881 CCGAGCCCGACATTCCTTAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCAGT 2940
Db 3579 CCGAGCCCGACATTCCTTAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCAGT 3638
QY 2941 GACACTGGCCCTCCAGGTGTCAGCCAGCAGCTGCCAGGCCACCTGGAGCAAGAGCG 3000
Db 3639 GACACTGGCCCTCCAGGTGTCAGCCAGCAGCTGCCAGGCCACCTGGAGCAAGAGCG 3698
QY 3001 A 3001
Db 3699 A 3699

RESULT 7
AX430856
LOCUS AX430856
DEFINITION Sequence 1 from Patent WO0240683.
ACCESSION AX430856
VERSION AX430856.1 GI:21655921
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Ketchum,K., Beagley,E.M., Wei,M.H. and di Francesco,V.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding
human kinase proteins, and uses thereof
JOURNAL Patent: WO 0240683-A 1 23-MAY-2002;
FEATURES
PE CORP NY (US)
source Location/Qualifiers
1..5207
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 99.9%; Score 2997.8; DB 6; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCTACGAGCAGAACTCGGGCCCGAGGATACAGGAGCAGACATCTCTGGCCGCGCTGAG 60
Db 699 CCTACGAGCAGAACTCGGGCCCGAGGATACAGGAGCAGACATCTCTGGCCGCGCTGAG 758
QY 61 CCACCCGCTGGTACGCGGGCTGTGGACCAAGTTTGGAGACCCCGCAGACCCCTCATCTCAT 120
```


Db 759 CACCCCGCTGGTCA CGGGGCTGTG GACCAAGTTTGAGACCGCAAGCCCTCATCTCAT 818
Qy 121 CTTGGAGCTGTGCTCA TCCGAGAGCTGTG GACCGCTGTACAGGAAGGGGTGTGAC 180
Db 819 CTTGGAGCTGTGCTCA TCCGAGAGCTGTG GACCGCTGTACAGGAAGGGGTGTGAC 878
Qy 181 GGAGGCGAGGTCAAGGT CTACATCCAGCAGCTGTG TGAGGGGCTGCACTACCTGCAACAG 240
Db 879 GGAGGCGAGGTCAAGGT CTACATCCAGCAGCTGTG TGAGGGGCTGCACTACCTGCAACAG 938
Qy 241 CCATGGGTCTTCCACT TGACATATAAGCCCTT AAACATCTGATGCTCATCTGTCGCG 300
Db 939 CCATGGGTCTTCCACT TGACATATAAGCCCTT AAACATCTGATGCTCATCTGTCGCG 998
Qy 301 GGAAGACATTAATAAT CTGGGACTTTGGCT TTGGCCAGAAACATCAACCCAGAGAGCTGCA 360
Db 999 GGAAGACATTAATAAT CTGGGACTTTGGCT TTGGCCAGAAACATCAACCCAGAGAGCTGCA 1059
Qy 361 GTTCAGCCAGTACGGCT CCCCCTGAGTTCTG TCTCCCGAGATCATCCAGCAGAACCTGT 420
Db 1059 GTTCAGCCAGTACGGCT CCCCCTGAGTTCTG TCTCCCGAGATCATCCAGCAGAACCTGT 1118
Qy 421 GAGCGAAGCTCCGACAT TTGGCCAGTGGTGTG TCACTCTACCTCAGCTGACCTGCTC 480
Db 1119 GAGCGAAGCTCCGACAT TTGGCCAGTGGTGTG TCACTCTACCTCAGCTGACCTGCTC 1178
Qy 481 ATCCCATTTGCGGCGAGGTG AACGCTGCTG CCACTCTGAAAGCTCTTGAGGGGCGGT 540
Db 1179 ATCCCATTTGCGGCGAGGTG AACGCTGCTG CCACTCTGAAAGCTCTTGAGGGGCGGT 1238
Qy 541 GTTCAGGAGCAGCCGAT TGGTGGCCACTCAG CCGAAGACGCCAAAGACTTTCATCAAGGC 600
Db 1239 GTTCAGGAGCAGCCGAT TGGTGGCCACTCAG CCGAAGACGCCAAAGACTTTCATCAAGGC 1298
Qy 601 TACGCTGACAGAGCCCT CAGGCGCGGCTAG TGGCGGCCAGTGCCTCTCCACCCCTG 660
Db 1299 TACGCTGACAGAGCCCT CAGGCGCGGCTAG TGGCGGCCAGTGCCTCTCCACCCCTG 1358
Qy 661 GTTCTGAAATCCATGCT TGGGAGAGGCCCCA CTTTCATCAACCAAGCAGCTCAAGTT 720
Db 1359 GTTCTGAAATCCATGCT TGGGAGAGGCCCCA CTTTCATCAACCAAGCAGCTCAAGTT 1418
Qy 721 CCTCTGCGCCGAGTCTG CAGGCTTCCGTGATG AGTACAGTCCATCTCTGGTGTAT 780
Db 1419 CCTCTGCGCCGAGTCTG CAGGCTTCCGTGATG AGTACAGTCCATCTCTGGTGTAT 1478
Qy 781 GCGCTCCATCCTGAGCTG CTGGGGGCCAC CCGACAGCCCTCCCTCGGCGTAGCCG 840
Db 1479 GCGCTCCATCCTGAGCTG CTGGGGGCCAC CCGACAGCCCTCCCTCGGCGTAGCCG 1538
Qy 841 GCACTCTGCAAGGACACT GTGTGCTCCTCAGT TCTCCTCTCTGACAAACAGCT 900
Db 1539 GCACTCTGCAAGGACACT GTGTGCTCCTCAGT TCTCCTCTCTGACAAACAGCT 1598
Qy 901 CGCCCCATTTCCCGGGTATG TCACTGCCAC CTTCCCGGTGACACTCAACACTGCT 960
Db 1599 CGCCCCATTTCCCGGGTATG TCACTGCCAC CTTCCCGGTGACACTCAACACTGCT 1658
Qy 961 GCACCCCGGGGCTTCTCT GCGGCCCTCGG CCGAGCTGCTGAGGAAGCCGAGGCCAGTGA 1020
Db 1659 GCACCCCGGGGCTTCTCT GCGGCCCTCGG CCGAGCTGCTGAGGAAGCCGAGGCCAGTGA 1718
Qy 1021 GCGTTCACCGAGGCCGAGT CCGCTTGATNTT CCGAGGGTGC GGGCCACCGGCGC 1080
Db 1719 GCGTTCACCGAGGCCGAGT CCGCTTGATNTT CCGAGGGTGC GGGCCACCGGCGC 1778
Qy 1081 CCAGGGTGTGTCGCCCGG CACAGCTCATCC GCGAGCTGTTCTTCAACACAGGGGTGA 1140
Db 1779 CCAGGGTGTGTCGCCCGG CACAGCTCATCC GCGAGCTGTTCTTCAACACAGGGGTGA 1838
Qy 1141 GAGCCCTGAGCA CGGGCCCTTGGCC CCGGGAGCAGGGCGCA CCGGCGCGGGCGCA 1200

Db 1839 GAGCCCTGAGCA CGGGGCTCTG GCGCGGGAGCAGGGGGCACCCGGGCCCGCGCGGCA 1898
Qy 1201 CTTGCTGAAGGCGGGGTAC ATTTGCGGGGCGCTG CCGAGGCTTGC CGAGCCACTGATGGA 1260
Db 1899 CTTGCTGAAGGCGGGGTAC ATTTGCGGGGCGCTG CCGAGGCTTGC CGAGCCACTGATGGA 1958
Qy 1261 GCACCGGCTGCTGAGAGGAGGCGCGC AGGAGGAGCAGGCCACCCCTCTCTGGGCCAAAGC 1320
Db 1959 GCACCGGCTGCTGAGAGGAGGCGCGC AGGAGGAGCAGGCCACCCCTCTCTGGGCCAAAGC 2018
Qy 1321 CCCCTCATTTGAGAGCTG CCGCTCCGCTG CCGCTCTGCGACCCCACTTGGCCCCCTGGCCA 1380
Db 2019 CCCCTCATTTGAGAGCTG CCGCTCCGCTG CCGCTCTGCGACCCCACTTGGCCCCCTGGCCA 2078
Qy 1381 CAGCCACTCCCTGGAACATGACTCTCCGAGGAC CCCCCCGCCCTCTCTCGAGGCGCTCGG 1440
Db 2079 CAGCCACTCCCTGGAACATGACTCTCCGAGGAC CCCCCCGCCCTCTCTCGAGGCGCTCGG 2138
Qy 1441 TGAGGCAACAGCACTGCTT CAGCCCTTCCG GGGGGGCCCCCTATCAGGGACATGGGGCA 1500
Db 2139 TGAGGCAACAGCACTGCTT CAGCCCTTCCG GGGGGGCCCCCTATCAGGGACATGGGGCA 2198
Qy 1501 CCCTCAGGGTCCAAAGCAGCTTCCATCCACTG TGTGGCCACCAGGCACTGCTCAGCCAGA 1560
Db 2199 CCCTCAGGGTCCAAAGCAGCTTCCATCCACTG TGTGGCCACCAGGCACTGCTCAGCCAGA 2258
Qy 1561 GAGGCCATCCCGGACAGCCCTTGGGGGCGAG CCGCCCTTCTGCGACCCCAAGCAGGG 1620
Db 2259 GAGGCCATCCCGGACAGCCCTTGGGGGCGAG CCGCCCTTCTGCGACCCCAAGCAGGG 2318
Qy 1621 TTCTGCCCCCAGGAGGGCTG CAGCCCCCA CCGCAGCAGTGTGCCCATGCCCTCTCTGGCTC 1680
Db 2319 TTCTGCCCCCAGGAGGGCTG CAGCCCCCA CCGCAGCAGTGTGCCCATGCCCTCTCTGGCTC 2378
Qy 1681 CTTTCCCTCAGAGATCTTGTCAAGAGGCGCC CTTAGTACCTCAAGCCCTTCTTTGGGACA 1740
Db 2379 CTTTCCCTCAGAGATCTTGTCAAGAGGCGCC CTTAGTACCTCAAGCCCTTCTTTGGGACA 2438
Qy 1741 GCCCCAGGCAACCCCTG CCGCTGCCAAAGCAG CCCCCCATTTGGACTCTAAGATGGGGCC 1800
Db 2439 GCCCCAGGCAACCCCTG CCGCTGCCAAAGCAG CCCCCCATTTGGACTCTAAGATGGGGCC 2498
Qy 1801 TGAGACATCTCTCTTCTCTGGAGGCGCAAAA CCGGCCCCCTGCGAGTTCCCAAGGGTCAGC 1860
Db 2499 TGAGACATCTCTCTTCTCTGGAGGCGCAAAA CCGGCCCCCTGCGAGTTCCCAAGGGTCAGC 2558
Qy 1861 CTTCCAGGCGAGCTCTTCCCAAGTAGCTCC CTTCAGGGTGGGCTCCTCCAGGTGGGCAC 1920
Db 2559 CTTCCAGGCGAGCTCTTCCCAAGTAGCTCC CTTCAGGGTGGGCTCCTCCAGGTGGGCAC 2618
Qy 1921 AGAGCTTGGCCCTCCCTG GATGCGAGGGCTG GACCCAGGAGGCTGAGGATCTGTCCGA 1980
Db 2619 AGAGCTTGGCCCTCCCTG GATGCGAGGGCTG GACCCAGGAGGCTGAGGATCTGTCCGA 2678
Qy 1981 CTTCCACAACCTTTGCA CGCCCTCAGGAA CAGGCGACCATGCGCAAGTTCTCTCCCTGGG 2040
Db 2679 CTTCCACAACCTTTGCA CGCCCTCAGGAA CAGGCGACCATGCGCAAGTTCTCTCCCTGGG 2738
Qy 2041 TGGTCCGGGGGCTACG CAGGGCTGGCTGGCTG ATGGCACTTTGGCTTTGGTGGAGATGC 2100
Db 2739 TGGTCCGGGGGCTACG CAGGGCTGGCTGGCTG ATGGCACTTTGGCTTTGGTGGAGATGC 2798
Qy 2101 AGGGGCGATGCTGGG GCGAGGGGCCCATGTG TGGGCCAGGATAGCTGGGCTGTGTCCAGTTC 2160
Db 2799 AGGGGCGATGCTGGG GCGAGGGGCCCATGTG TGGGCCAGGATAGCTGGGCTGTGTCCAGTTC 2858
Qy 2161 GAGAGAGAGGAGCAGGAG GCGCAGGGCTGAGT CCCCAGTCCGAGGAGCAGCAGAGGCG 2220
Db 2859 GAGAGAGAGGAGCAGGAG GCGCAGGGCTGAGT CCCCAGTCCGAGGAGCAGCAGAGGCG 2918
Qy 2221 CAGGGCTCAGAGCCCACTG CCGCCCAAGGT CAGTGCAGGGCTGTGTCTGAGGTGGCAGGGC 2280
Db 2919 CAGGGCTCAGAGCCCACTG CCGCCCAAGGT CAGTGCAGGGCTGTGTCTGAGGTGGCAGGGC 2978


```
QY 2281 TCCACACAGGAGCTCTCCAGAGCCCAACCCCATGGAGAGCATCGGGCAGGTCTCCCTGGT 2340
DB 2979 TCCACACAGGAGCTCTCCAGAGCCCAACCCCATGGAGAGCATCGGGCAGGTCTCCCTGGT 3038
QY 2341 CGAGATCCGGGACCTGTCCAGGTGATCGGAGGCGCCGACACAAATATCCCTGGACATTTTC 2400
DB 3039 GCAGATCCGGGACCTGTCCAGGTGATCGGAGGCGCCGACACAAATATCCCTGGACATTTTC 3098
QY 2401 CGAGTGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
DB 3099 CGAGTGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 3158
QY 2461 CGAGTTTATGATCTTCAGAAAGTCCCAAGTCCCTCAGCAGAGCGCCCTCCCCCAT 2520
DB 3159 CGAGTTTATGATCTTCAGAAAGTCCCAAGTCCCTCAGCAGAGCGCCCTCCCCCAT 3218
QY 2521 GCCTGAGGAGGAGCTGGCGGAGTCCCGAGGCGCACGTCGGCCCTGGCAGGTGAACCTGGG 2580
DB 3219 GCCTGAGGAGGAGCTGGCGGAGTCCCGAGGCGCACGTCGGCCCTGGCAGGTGAACCTGGG 3278
QY 2591 CCCCCACGAGCCCTGGAGATCACAGAGGAGTCAAGAGATGTGACGCGCTGCTGGCAGA 2640
DB 3279 CCCCCACGAGCCCTGGAGATCACAGAGGAGTCAAGAGATGTGACGCGCTGCTGGCAGA 3338
QY 2641 GGCTGCCGTGGCAGGAGCGCAAGTGTCTCGCCGCTCAACGAGCTCTTCCACTTCCC 2700
DB 3339 GGCTGCCGTGGCAGGAGCGCAAGTGTCTCGCCGCTCAACGAGCTCTTCCACTTCCC 3398
QY 2701 TGGGAGGACCTTCGCGCTGGATGAGCTTCAGAGCTGGGGCTGCTGAGAGATCAAGGC 2760
DB 3399 TGGGAGGACCTTCGCGCTGGATGAGCTTCAGAGCTGGGGCTGCTGAGAGATCAAGGC 3458
QY 2761 CTCCTGGAGACATCTCCCGATCCTGAAGGCGAGCGCGAAGTCTGGAGAGAGGG 2820
DB 3459 CTCCTGGAGACATCTCCCGATCCTGAAGGCGAGCGCGAAGTCTGGAGAGAGGG 3518
QY 2821 GCGGCGGAGGAGAGAGCAGGCTGCTCTTCCTCGGCTCTCAGGTCTGAGAGCTGGGA 2880
DB 3519 GCGGCGGAGGAGAGAGCAGGCTGCTCTTCCTTCGCGCTCTCAGGTCTGAGAGCTGGGA 3578
QY 2881 CCGAGCGCGGACATCTTAAGGAGCTCTCAGATGAGAGTGTGCTGGGCGAGTCACT 2940
DB 3579 CCGAGCGCGGACATCTTAAGGAGCTCTCAGATGAGAGTGTGCTGGGCGAGTCACT 3638
QY 2941 GACACTGGCCTGCCAGGTGTGAGCCAGCAGCTGCCAGGCGCACTGGAGCAAGAGCGG 3000
DB 3639 GACACTGGCCTGCCAGGTGTGAGCCAGCAGCTGCCAGGCGCACTGGAGCAAGAGCGG 3698
QY 3001 A 3001
DB 3699 A 3699

RESULT 8
AX430858
LOCUS AX430858 5207 bp DNA linear PAT 28-JUN-2002
DEFINITION Sequence 3 from Patent WO0240683.
ACCESSION AX430858
VERSION AX430858.1 GI:21655922
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Ketchum, K., Beasley, E.M., Wei, M.H. and di Francesco, V.
AUTHORS Isolated human kinase proteins, nucleic acid molecules encoding
TITLE human kinase proteins, and uses thereof
JOURNAL Patent: WO 0240683-A 3 23-MAY-2002;
PE CORP NY (US)
FEATURES
Location/Qualifiers
source 1..5207
```

```
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

ORIGIN

```
Query Match 99.9%; Score 2997.8; DB 6; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTACGGAGCAGAACTCGGGCCCGAGGCATACAGGAGCGAGACATCTCTGGCGCGCTGAG 60
DB 699 CCTACGGAGCAGAACTCGGGCCCGAGGCATACAGGAGCGAGACATCTCTGGCGCGCTGAG 758
QY 61 CCACCCGCTGTGTACGGGGCTGTGGACCAAGTTTGAACCCGCAAGACCTCTCATCTCAT 120
DB 759 CCACCCGCTGTGTACGGGGCTGTGGACCAAGTTTGAACCCGCAAGACCTCTCATCTCAT 818
QY 121 CCTGAGAGCTGTCTCATCCGAGGAGCTGTCTGACCGCTGTACAGGAAGGCGGTGTGAC 180
DB 819 CCTGAGAGCTGTCTCATCCGAGGAGCTGTCTGACCGCTGTACAGGAAGGCGGTGTGAC 878
QY 181 GGAGCCGAGGTCAAGGTCTTACATCCAGAGCTGTGTGGAGGGGCTGCACTACCTGCACAG 240
DB 879 GGAGCCGAGGTCAAGGTCTTACATCCAGAGCTGTGTGGAGGGGCTGCACTACCTGCACAG 938
QY 241 CCATGGCGTTCTCCACCTGTGACATAAAGCCCTCTAAACATCTCTGATGTGTGATCTCTGCCG 300
DB 939 CCATGGCGTTCTCCACCTGTGACATAAAGCCCTCTAAACATCTCTGATGTGTGATCTCTGCCG 998
QY 301 GGAAGACATTAATAATCTCGGACTTTGGCTTTGGCCAGAACATCAACCCAGAGAGCTGCA 360
DB 999 GGAAGACATTAATAATCTCGGACTTTGGCTTTGGCCAGAACATCAACCCAGAGAGCTGCA 1058
QY 361 GTTCAGCGAGTACCGGCTCCCTGAGTTCTCTCCCGAGAGATCATCCAGAGAGAACTCTGT 420
DB 1059 GTTCAGCGAGTACCGGCTCCCTGAGTTCTCTCCCGAGAGATCATCCAGAGAGAACTCTGT 1118
QY 421 GAGGAGAGCTTCCGACATTTTGGGCGATGGGTGTGTATCTCTACCTCAGCCTGACCTGTCTC 480
DB 1119 GAGGAGAGCTTCCGACATTTTGGGCGATGGGTGTGTATCTCTACCTCAGCCTGACCTGTCTC 1178
QY 481 ATCCCCATTTCCCGCGAGAGTGACCGTGCACACCTCTCTGAAACGTCCTCGAGGGGCGGT 540
DB 1179 ATCCCCATTTCCCGCGAGAGTGACCGTGCACACCTCTCTGAAACGTCCTCGAGGGGCGGT 1238
QY 541 GTCATGAGCAGAGCCCATGCTGCTGCCACCTCAGGAGAGACGCCAAAGACTTTCATCAAGGC 600
DB 1239 GTCATGAGCAGAGCCCATGCTGCTGCCACCTCAGGAGAGACGCCAAAGACTTTCATCAAGGC 1298
QY 601 TACGCTGCAGAGAGCCCTCAGGCGCGGCTTAGTGTGGGCGCCAGTGTCTCTCCACCCCTG 660
DB 1299 TACGCTGCAGAGAGCCCTCAGGCGCGGCTTAGTGTGGGCGCCAGTGTCTCTCCACCCCTG 1358
QY 661 GTTCCTGAATTCATGCTCGGAGGAGGCCACTTCATCAACCAAGAGAGCTCAAGTT 720
DB 1359 GTTCCTGAATTCATGCTCGGAGGAGGCCACTTCATCAACCAAGAGAGCTCAAGTT 1418
QY 721 CCTCTGCGCCGAAAGTGTCTGCGCAGCGTTCCTGATGAGCTACAGTCCATCTCTGGTGTAT 780
DB 1419 CCTCTGCGCCGAAAGTGTCTGCGCAGCGTTCCTGATGAGCTACAGTCCATCTCTGGTGTAT 1478
QY 781 GCGTTCATCTCCTGAGTGTGTGGGGGCGCAACCCGACAGCCCTCTCTCTGGGCTAGCCCG 840
DB 1479 GCGTTCATCTCCTGAGTGTGTGGGGGCGCAACCCGACAGCCCTCTCTCTGGGCTAGCCCG 1538
QY 841 GCACCTCTGAGGAGACACTGTGTGCTCTCCAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
DB 1539 GCACCTCTGAGGAGACACTGTGTGCTCTCCAGTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1598
QY 901 CGCCCCATTTCCCGGGCTAAGTCACTGCCACCTCTCCCGGTGACACACTCACCACCTGTCT 960
DB 1599 CGCCCCATTTCCCGGGCTAAGTCACTGCCACCTCTCCCGGTGACACACTCACCACCTGTCT 1658
```


QY 961 GCACCCCGGGGCTTCTCGGGCCCTCGGCAGCCTCGCTGAGGAAGCCAGGCCAGTGA 1020
| | | | |
Db 1659 GCACCCCGGGGCTTCTCGGGCCCTCGGCAGCCTCGCTGAGGAAGCCAGGCCAGTGA 1718
| | | | |
QY 1021 GGCCTCCACGAGGCCCTCGAGCTCGGCTGATCTCCGAGGGTGCAGGGCCACCGGCCGC 1080
| | | | |
Db 1719 GGCCTCCACGAGGCCCTCGAGCTCGGCTGATCTCCGAGGGTGCAGGGCCACCGGCCGC 1778
| | | | |
QY 1081 CCAGGGCTGCTGCTCCCGGGCAGACAGCGTCATCCGACGCTGTTCATCCACAGCGGGGTGA 1140
| | | | |
Db 1779 CCAGGGCTGCTGCTCCCGGGCAGACAGCGTCATCCGACGCTGTTCATCCACAGCGGGGTGA 1838
| | | | |
QY 1141 GAGCCCTGAGCAGCGGGCCCTTGCCCGGGGAGCAGGGCCGACCCCGGCCCGGCCGCGCA 1200
| | | | |
Db 1839 GAGCCCTGAGCAGCGGGCCCTTGCCCGGGGAGCAGGGCCGACCCCGGCCCGGCCGCGCA 1898
| | | | |
QY 1201 CTGTCTGAAGGGGGGTCAATTGCGGGGGCGTGCAGGCTGCGCGAGCCACTGATGGA 1260
| | | | |
Db 1899 CTGTCTGAAGGGGGGTCAATTGCGGGGGCGTGCAGGCTGCGCGAGCCACTGATGGA 1958
| | | | |
QY 1261 GCACCGCGTGTGAGGAGGAGCGCGCCAGGAGGAGCAGGCGCACCTCTCTGGGCCAAAGC 1320
| | | | |
Db 1959 GCACCGCGTGTGAGGAGGAGCGCGCCAGGAGGAGCAGGCGCACCTCTCTGGGCCAAAGC 2018
| | | | |
QY 1321 CCCCTCATTCGAGACTGCCCCCTCGGCTGCTGCTGCGCAGCCCACTTGGGCCCTGGCCA 1380
| | | | |
Db 2019 CCCCTCATTCGAGACTGCCCCCTCGGCTGCTGCTGCGCAGCCCACTTGGGCCCTGGCCA 2078
| | | | |
QY 1381 CAGCCACTCCCTTGAACATGACTCTCCGAGCAGCCCCCGGCCCTCTCGGAGGCTGCGG 1440
| | | | |
Db 2079 CAGCCACTCCCTTGAACATGACTCTCCGAGCAGCCCCCGGCCCTCTCGGAGGCTGCGG 2138
| | | | |
QY 1441 TGAGGACAGCGACTGCTTCCAGCCCTTCGCGGGGGGCGCCCTATTCAGGGACATGGGGCA 1500
| | | | |
Db 2139 TGAGGACAGCGACTGCTTCCAGCCCTTCGCGGGGGGCGCCCTATTCAGGGACATGGGGCA 2198
| | | | |
QY 1501 CCTCAGGGCTCCAGAGGCTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGA 1560
| | | | |
Db 2199 CCTCAGGGCTCCAGAGGCTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGA 2258
| | | | |
QY 1561 GAGGCCATCCCGGACAGCCCTTGGGGGCGAGCAGCCCTTCTGCGCACCCCAAGCAGGG 1620
| | | | |
Db 2259 GAGGCCATCCCGGACAGCCCTTGGGGGCGAGCAGCCCTTCTGCGCACCCCAAGCAGGG 2318
| | | | |
QY 1621 TTCTGCCCCCAGAGGGCTGAGCCCCCAACCGAGTGGCCCACTGCTCTGCTC 1680
| | | | |
Db 2319 TTCTGCCCCCAGAGGGCTGAGCCCCCAACCGAGTGGCCCACTGCTCTGCTC 2378
| | | | |
QY 1681 CTTCCTCCAGGATCTTGCAGAGGGCCCTTAGTACCTCAAGCCCTCTTGGGACA 1740
| | | | |
Db 2379 CTTCCTCCAGGATCTTGCAGAGGGCCCTTAGTACCTCAAGCCCTCTTGGGACA 2438
| | | | |
QY 1741 GCCCAGGCACCCCCTGCCCCGCAAGAGCCGCCCAATTTGACTCTAAGATGGGGCC 1800
| | | | |
Db 2439 GCCCAGGCACCCCCTGCCCCGCAAGAGCCGCCCAATTTGACTCTAAGATGGGGCC 2498
| | | | |
QY 1801 TGGAGACATCTCTTCTTGGAGGCCAAACCCGGGCCCTGCGATTTCCCGAGGGTCAGC 1860
| | | | |
Db 2499 TGGAGACATCTCTTCTTGGAGGCCAAACCCGGGCCCTGCGATTTCCCGAGGGTCAGC 2558
| | | | |
QY 1861 CTCACGCGAGCTCTTCCCAAGTGAGCTCCCTCAGGGTGGGTCTCTCCAGAGTGGGCAC 1920
| | | | |
Db 2559 CTCACGCGAGCTCTTCCCAAGTGAGCTCCCTCAGGGTGGGTCTCTCCAGAGTGGGCAC 2618
| | | | |
QY 1921 AGAGCTGGCCCCCTCTCGATGCGAGGGGTGGACCCAGAGGCTGAGGATCTGTCCGA 1980
| | | | |
Db 2619 AGAGCTGGCCCCCTCTCGATGCGAGGGGTGGACCCAGAGGCTGAGGATCTGTCCGA 2678
| | | | |
QY 1981 CTCACACCCACCTTGCAGGGGCTCAGGAACAGGGCCACCATGGGCAAGTTCTCCCTGGG 2040
| | | | |
Db 2679 CTCACACCCACCTTGCAGGGGCTCAGGAACAGGGTACCATGGGCAAGTTCTCCCTGGG 2738
| | | | |
QY 2041 TGGTTCGGGGGGCTACGACGGCGTGGCTGCTATGGCACCTTTTGCTCTTGGTGGAGATGC 2100
| | | | |

Db 2739 TGGTTCGGGGGGCTACGACGCGTGGCTATGGCACCTTTTGCTTTGGTGGAGATGC 2798
| | | | |
QY 2101 AGGGGGCATGCTGGGGCAGGGGCCCATGTGGGGCAGGATAGCCTGGGCTGTGTCCCCAGTC 2160
| | | | |
Db 2799 AGGGGGCATGCTGGGGCAGGGGCCCATGTGGGGCAGGATAGCCTGGGCTGTGTCCCCAGTC 2858
| | | | |
QY 2161 GGAGGAGGAGGAGCAGGAGGAGGCCAGGGCTGAGTCCAGTTCGAGAGCAGCAGGAGGC 2220
| | | | |
Db 2859 GGAGGAGGAGGAGCAGGAGGAGGCCAGGGCTGAGTCCAGTTCGAGAGCAGCAGGAGGC 2918
| | | | |
QY 2221 CAGGGCTGAGAGGCCCATCTGCCCCCAGGTGAGTGCAGGTTCAGGTTCGCGCAGGCG 2280
| | | | |
Db 2919 CAGGGCTGAGAGGCCCATCTGCCCCCAGGTGAGTGCAGGTTCGCGCAGGCG 2978
| | | | |
QY 2281 TCCACACAGAGGCTCTCCAGAGGCCACCCCATGGGAGGACATCGGGCAGGTCTCCCTGGT 2340
| | | | |
Db 2979 TCCACACAGAGGCTCTCCAGAGGCCACCCCATGGGAGGACATCGGGCAGGTCTCCCTGGT 3038
| | | | |
QY 2341 GCAGATCCGGGACCTGTGAGTGCAGGAGGCCGACCAATATCCCTGGACATTTTC 2400
| | | | |
Db 3039 GCAGATCCGGGACCTGTGAGTGCAGGAGGCCGACCAATATCCCTGGACATTTTC 3098
| | | | |
QY 2401 CGAGGTGAGCCCGCCCTTACCTCAACCTCTCAGACCTGTGACGATATCAAGTACCTCCCATTT 2460
| | | | |
Db 3099 CGAGGTGAGCCCGCCCTTACCTCAACCTCTCAGACCTGTGACGATATCAAGTACCTCCCATTT 3158
| | | | |
QY 2461 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCGCCCTCCCCCAT 2520
| | | | |
Db 3159 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCGCCCTCCCCCAT 3218
| | | | |
QY 2521 GGCTGAGGAGGAGCTGCGCCGAGTTCCGGAGCCACAGTGGCCCTGCGCAGGTGAAGTGGG 2580
| | | | |
Db 3219 GGCTGAGGAGGAGCTGCGCCGAGTTCCGGAGCCACAGTGGCCCTGCGCAGGTGAAGTGGG 3278
| | | | |
QY 2581 CCCCCACGAGGCTGGAGATCAGAGGAGTCAAGAGGATGTGGACCGCTGTCTGGCAGA 2640
| | | | |
Db 3279 CCCCCACGAGGCTGGAGATCAGAGGAGTCAAGAGGATGTGGACCGCTGTCTGGCAGA 3338
| | | | |
QY 2641 GGCTGCGCTGGGCGAGGAGCGCAAGTGTCTCGCCGTCAGCGAGCTCTTCCACTTCCC 2700
| | | | |
Db 3339 GGCTGCGCTGGGCGAGGAGCGCAAGTGTCTCGCCGTCAGCGAGCTCTTCCACTTCCC 3398
| | | | |
QY 2701 TGGAGGCACTCTGCGCTGGATGAGCTTCAGAGCTGGGGCTGCGTGAGAGGTGAAGGC 2760
| | | | |
Db 3399 TGGAGGCACTCTGCGCTGGATGAGCTTCAGAGCTGGGGCTGCGTGAGAGGTGAAGGC 3458
| | | | |
QY 2761 CTCGTTGGAGCACATCTCCGGATCTGAAGGGCAGGCCGGAAGGTCTGGAGAGGAGGC 2820
| | | | |
Db 3459 CTCGTTGGAGCACATCTCCGGATCTGAAGGGCAGGCCGGAAGGTCTGGAGAGGAGGC 3518
| | | | |
QY 2821 GCCCCCAGGAGAGAGCCAGGCTTCTTCTTCCGGCTCTCAGGTCTCAAGAGCTGGGA 2880
| | | | |
Db 3519 GCCCCCAGGAGAGAGCCAGGCTTCTTCTTCCGGCTCTCAGGTCTCAAGAGCTGGGA 3578
| | | | |
QY 2881 CCGAGCGCCGACATTCCTTAAGGAGCTCTCAGATGAGACTGTGGTCTCTGGGCGAGTCACT 2940
| | | | |
Db 3579 CCGAGCGCCGACATTCCTTAAGGAGCTCTCAGATGAGACTGTGGTCTCTGGGCGAGTCACT 3638
| | | | |
QY 2941 GACACTGGGCTGCCAGGTGTGAGCCAGCCAGCTGCCAGGCCACCTGGAGCAAGAAGC 3000
| | | | |
Db 3639 GACACTGGGCTGCCAGGTGTGAGCCAGCCAGCTGCCAGGCCACCTGGAGCAAGAAGC 3698
| | | | |
QY 3001 A 3001
|
Db 3699 A 3699

RESULT 9
AX504255
LOCUS
DEFINITION Sequence 44 from Patent WO0233099.
AX504255
ACCESSION
linear PAT 27-SEP-2002

AX504255.1 GI:23386095
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Gururajan, R., Baughn, M.R., Wallia, N.K., Elliott, V.S., Xu, Y.,
Arvizu, C., Yeo, M.G., Ramkumar, J., Ding, L., Tang, Y.T., Hafalia, A.J.,
Nguyen, D.B., Gandhi, A.R., Lu, Y., Yue, H., Burford, N., Bandman, O.,
Tribouley, C.M., Lal, P.G., Recipon, S.A., Lu, D.A., Borowsky, M.L.,
Thornton, M., Swarnaker, A., Thangavelu, K., Khan, F.A. and Ison, C.H.
Human kinases
TITLE Patent: WO 0233099-A 44 25-APR-2002;
JOURNAL Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1..5454
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 7638121CB1"
ORIGIN
Query Match 99.8%; Score 2996.2; DB 6; Length 5454;
Best Local Similarity 99.9%; Pred No. 0;
Matches 2998; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
Qy 1 CTAAGGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTTGGCGCGCTGAG 60
Db CTAAGGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTTGGCGCGCTGAG 1022
Qy 61 CCACCCGCTGTGTCACGGGGCTGTGGACAGTGTGAGACCCGACAGACCTCATCTCAT 120
Db CCACCCGCTGTGTCACGGGGCTGTGGACAGTGTGAGACCCGACAGACCTCATCTCAT 1082
Qy 121 CTTGAGAGTGTCTCATCGAGGAGCTGTGGACCGCTGTACAGAAAGGGCGTGGTAC 180
Db CTTGAGAGTGTCTCATCGAGGAGCTGTGGACCGCTGTACAGAAAGGGCGTGGTAC 1142
Qy 181 GGAGCCGAGGTCAAGGTCTATCCAGAGCTGTGTGAGGGGTGCACTACCTGCAAG 240
Db GGAGCCGAGGTCAAGGTCTATCCAGAGCTGTGTGAGGGGTGCACTACCTGCAAG 1202
Qy 241 CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAAACATCTGTATGGTGCATCTGCGCG 300
Db CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAAACATCTGTATGGTGCATCTGCGCG 1262
Qy 301 GGAAGACATTAAATCTCGACTTTGGCTTTGGCCAGAAACATCAACCAGAGAGCTGCA 360
Db GGAAGACATTAAATCTCGACTTTGGCTTTGGCCAGAAACATCAACCAGAGAGCTGCA 1322
Qy 361 GTTCAGCCAGTACGGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGCAGAACCTGT 420
Db GTTCAGCCAGTACGGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGCAGAACCTGT 1382
Qy 421 GAGCGAAGCTCCGACATTTGGGCCATGGGTGTGATCTCTACCTCAGCCTGACCTGCTC 480
Db GAGCGAAGCTCCGACATTTGGGCCATGGGTGTGATCTCTACCTCAGCCTGACCTGCTC 1442
Qy 481 ATCCCCATTTGCGGCGAGAGTGACGTGCGACCCCTCTGACGCTCTGGAGGGGCGGT 540
Db ATCCCCATTTGCGGCGAGAGTGACGTGCGACCCCTCTGACGCTCTGGAGGGGCGGT 1502
Qy 541 GTTCATGAGCAGCCCCCATGGTGGCCACTCAGCGAAGACGCGCAAGACTTTCATCAAGGC 600
Db GTTCATGAGCAGCCCCCATGGTGGCCACTCAGCGAAGACGCGCGCAAGACTTTCATCAAGGC 1562
Qy 601 TACGCTCAGAGAGCCCTCAGGCCCGGCTAGTGGCGCCAGTGCCTCTCCACCCCTG 660
Db TACGCTCAGAGAGCCCTCAGGCCCGGCTAGTGGCGCCAGTGCCTCTCCACCCCTG 1622
Qy 661 GTTCTGAAATCCATGCTGGGAGGCGCCACTTTCATCAACCAAGCAGCTCAAGTT 720
Db GTTCTGAAATCCATGCTGGGAGGCGCCACTTTCATCAACCAAGCAGCTCAAGTT 2762

Db 1623 GTTCTGAAATCCATGCTGGGAGGCGCCACTTTCATCAACCAAGCAGCTCAAGTT 1682
Qy 721 CCTCTGGCCGGAAGTGTCTGGCAGCGTTTCTCTGATGAGCTACAAAGTCCATCTGGTGAT 780
Db 1683 CCTCTGGCCGGAAGTGTCTGGCAGCGTTTCTCTGATGAGCTACAAAGTCCATCTGGTGAT 1742
Qy 781 GCGTCTCATCTCTGAGTGTCTGGGGCCACCGACAGCCCTCCCTCGGCGTAGCCCG 840
Db 1743 GCGTCTCATCTCTGAGTGTCTGGGGCCACCGACAGCCCTCCCTCGGCGTAGCCCG 1802
Qy 841 GCACCTCTGCAAGGACACACTGTGTGCTCTCTCAGTGTCTCTCTCTCTCTGCAAGAGCT 900
Db 1803 GCACCTCTGCAAGGACACACTGTGTGCTCTCTCAGTGTCTCTCTCTCTCTGCAAGAGCT 1862
Qy 901 CGCCGCCATTTGCCCGGGCTAAGTCACTGCGACACCTCCCGGTGACACACTCACCACCTCT 960
Db 1863 CGCCGCCATTTGCCCGGGCTAAGTCACTGCGACACCTCCCGGTGACACACTCACCACCTCT 1922
Qy 961 GCACCCCGGGGCTTCTTGGGGCTCTCGGGCAGCTGTGCTGAGGAAGCCGAGGCGCAGTGA 1020
Db 1923 GCACCCCGGGGCTTCTTGGGGCTCTCGGGCAGCTGTGCTGAGGAAGCCGAGGCGCAGTGA 1982
Qy 1021 GCGCTTCCACGAGGCGCCAGCTCCGCTGTGATCTCCGAGGGTGC CGGGCCACCGGGCGCG 1080
Db 1983 GCGCTTCCACGAGGCGCCAGCTCCGCTGTGATCTCCGAGGGTGC CGGGCCACCGGGCGCG 2042
Qy 1081 CCAGGGTGTCTGCGGGCAGCGTCACTCCGAGAGCTGTCTACACAGGCGGGTGA 1140
Db 2043 CCAGGGTGTCTGCGGGCAGCGTCACTCCGAGAGCTGTCTACACAGGCGGGTGA 2102
Qy 1141 GAGCCCTGAGCAGCGGGCCCTGGCGGGGAGCAGCGGCACCGGGCCCGGGCGCGCA 1200
Db 2103 GAGCCCTGAGCAGCGGGCCCTGGCGGGGAGCAGCGGCACCGGGCCCGGGCGCGCA 2162
Qy 1201 CCTCTGAAGGCGGGTACATTGCGGGGGCTCTGCGAGGCTGCGCGAGCACACTGATGGA 1260
Db 2163 CCTCTGAAGGCGGGTACATTGCGGGGGCTCTGCGAGGCTGCGCGAGCACACTGATGGA 2222
Qy 1261 GCACCGGTGTGAGAGAGAGGGCGCGCAGGGAGAGCAGCGCACCTCTCTGGGCCAAGC 1320
Db 2223 GCACCGGTGTGAGAGAGAGGGCGCGCAGGGAGAGCAGCGCACCTCTCTGGGCCAAGC 2282
Qy 1321 CCCCTATTCCAGACTGCCCCCTCCGCTGCTCTGGCACCCACTTTGGGCCCTTGCGCA 1380
Db 2283 CCCCTATTCCAGACTGCCCCCTCCGCTGCTCTGGCACCCACTTTGGGCCCTTGCGCA 2342
Qy 1381 CAGCACTCTCCCTGGAACATGACTCTCCGAGCAGCCCCCGGCCCTCTCTCGAGGCGCTGCGG 1440
Db 2343 CAGCACTCTCCCTGGAACATGACTCTCCGAGCAGCCCCCGGCCCTCTCTCGAGGCGCTGCGG 2402
Qy 1441 TGAGGCAAGGCACTGCGCTTCAGCCCCCTCGGGGGGGCCCCCTATTCAGGAGCATGGGGCA 1500
Db 2403 TGAGGCAAGGCACTGCGCTTCAGCCCCCTCGGGGGGGCCCCCTATTCAGGAGCATGGGGCA 2462
Qy 1501 CCCTCAGGGCTCCAGCAGAGTTCCTACCTGTGGCCACCCAGGAGCAGCTGCTCAGCCAGA 1560
Db 2463 CCCTCAGGGCTCCAGCAGAGTTCCTACCTGTGGCCACCCAGGAGCAGCTGCTCAGCCAGA 2522
Qy 1561 GAGGCCATTCGCCGACAGCCCTTGGGGGAGCAGCCAGCCCTTTCTGCGCACCCCAAGCAGGG 1620
Db 2523 GAGGCCATTCGCCGACAGCCCTTGGGGGAGCAGCCCTTTCTGCGCACCCCAAGCAGGG 2582
Qy 1621 TTCTGCCCGCCAGAGGGGTGACGCCCGCCACCGCAGAGTTCGCCCATGCCCTCTCTGGCTC 1680
Db 2583 TTCTGCCCGCCAGAGGGGTGACGCCCGCCACCGCAGAGTTCGCCCATGCCCTCTCTGGCTC 2642
Qy 1681 CTTCCCTCCAGGATCTTGCAGAGGGCCCCCTTAGTACCTCTCAGCCCTCTTGGGACA 1740
Db 2643 CTTCCCTCCAGGATCTTGCAGAGGGCCCCCTTAGTACCTCTCAGCCCTCTTGGGACA 2702
Qy 1741 GCCCAGGAGCAGCCCTCTGCCCTCTGCGCAAGAGCCCCCACTTGGACTCTAAGATGGGGCC 1800
Db 2703 GCCCAGGAGCAGCCCTCTGCCCTCTGCGCAAGAGCCCCCACTTGGACTCTAAGATGGGGCC 2762


```
QY 1801 TGGAGACATCTCTCTTCTGGAGGCGCAAAACCCGGCCCTCTGCAGTTCCCCAGGFTCAGC 1860
Db 2763 TGGAGACATCTCTCTTCTGGAGGCGCAAAACCCGGCCCTCTGCAGTTCCCCAGGFTCAGC 2822
QY 1861 CTCCAGGCGAGCTCTTCCAAAGTAGCTCTCTCAGGTGGCTCTCTCCAGAGTGGGCAC 1920
Db 2823 CTCCAGGCGAGCTCTTCCAAAGTAGCTCTCTCAGGTGGCTCTCTCCAGAGTGGGCAC 2882
QY 1921 AGAGCTGGCCCTCTCTGATGCGAGGCTGGAACCCAGAGGCTGAGGATCTGTCCGA 1980
Db 2883 AGAGCTGGCCCTCTCTGATGCGAGGCTGGAACCCAGAGGCTGAGGATCTGTCCGA 2942
QY 1981 CTCCACACCCACTTGCAGCGGCTCAGGAACAGCGACCATGCGCAAGTTCTCCCTGGG 2040
Db 2943 CTCCACACCCACTTGCAGCGGCTCAGGAACAGGTGACCATGCGCAAGTTCTCCCTGGG 3002
QY 2041 TGGTCCGGGGGTACGACAGGCTGGCTATGGCACCTTTGGCTTTGGTGGAGATGC 2100
Db 3003 TGGTCCGGGGGTACGACAGGCTGGCTATGGCACCTTTGGCTTTGGTGGAGATGC 3062
QY 2101 AGGGGCGATGCTGGGCGAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGCCAGTC 2160
Db 3063 AGGGGCGATGCTGGGCGAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGCCAGTC 3122
QY 2161 GGAGGAGGAGGACGAGGAGGCGGCGAGGCTGAGTCCAGTGGAGGAGCAGCAGGAGGC 2220
Db 3123 GGAGGAGGAGGACGAGGAGGCGGCGAGGCTGAGTCCAGTGGAGGAGCAGCAGGAGGC 3182
QY 2221 CAGGCTGAGAGCCACTGCCGCCAGCTCAGTGCAGAGGCTGTGCTGAGGTGGGAGGCG 2280
Db 3183 CAGGCTGAGAGCCACTGCCGCCAGCTCAGTGCAGAGGCTGTGCTGAGGTGGGAGGCG 3242
QY 2281 TCCACACAGGAGCTCTCAGAGGCCACCCCATGGGAGGACATCGGCGAGGCTCCCTGGT 2340
Db 3243 TCCACACAGGAGCTCTCAGAGGCCACCCCATGGGAGGACATCGGCGAGGCTCCCTGGT 3302
QY 2341 CGAGATCCGGGACCTGTGAGTGATGCGGAGCGCGCGACCAATATCTCTGGACATTTTC 2400
Db 3303 CGAGATCCGGGACCTGTGAGTGATGCGGAGCGCGCGACCAATATCTCTGGACATTTTC 3362
QY 2401 CGAGTGACCCCGCTACCTCACTCTCAGACTGACGATGACGATCAAGTACCTCCCAT 2460
Db 3363 CGAGTGACCCCGCTACCTCACTCTCAGACTGACGATGACGATCAAGTACCTCCCAT 3422
QY 2461 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGCGCGCTCCCCAT 2520
Db 3423 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGCGCGCTCCCCAT 3482
QY 2521 GGCTGAGGAGAGCTGGCCGAGTTCCCGGAGCCCAAGTGGCCCTGGCCAGGTGAATGGG 2580
Db 3483 GGCTGAGGAGAGCTGGCCGAGTTCCCGGAGCCCAAGTGGCCCTGGCCAGGTGAATGGG 3542
QY 2581 CCCACAGCGCTCGAGATCACAGAGGAGTCAAGAGATGTGACCGGCTGTGGCAGAG 2640
Db 3543 CCCACAGCGCTCGAGATCACAGAGGAGTCAAGAGATGTGACCGGCTGTGGCAGAG 3602
QY 2641 GGCTGCCGTGGCGAGGAGCGCAAGTGGTCTCCCGCTCAGCGAGCGCTCTTCCACTTCCC 2700
Db 3603 GGCTGCCGTGGCGAGGAGCGCAAGTGGTCTCCCGCTCAGCGAGCGCTCTTCCACTTCCC 3662
QY 2701 TGGGAGCAGCTGCGGCTGGATGAGCTTCAGAGCTGGGCTGGGTGAGAGGTGAAGGC 2760
Db 3663 TGGGAGCAGCTGCGGCTGGATGAGCTTCAGAGCTGGGCTGGGTGAGAGGTGAAGGC 3722
QY 2761 CTCGCTGGAGCACATCTCCCGGATCTGAAGGGCAGGCCGGAAGTCTGAGAGAGGAGG 2820
Db 3723 CTCGCTGGAGCACATCTCCCGGATCTGAAGGGCAGGCCGGAAGTCTGAGAGAGGAGG 3782
QY 2821 GCGGGCAGGAGAGGAGCGGCTTGTCTCTTCCGCTCTCAGGTCTGAAGAGCTGGGA 2880
Db 3783 GCGGGCAGGAGAGGAGCGGCTTGTCTCTTCCGCTCTCAGGTCTGAAGAGCTGGGA 3842
```

```
QY 2881 CCGAGCCGCGACATTCCTTAGGGAGCTCTCAGATGAGACTGTGGTCTCTGGGCCAGTCAGT 2940
Db 3843 CCGAGCCGCGACATTCCTTAGGGAGCTCTCAGATGAGACTGTGGTCTCTGGGCCAGTCAGT 3902
QY 2941 GACACTGGCTGCGCAGGTGTGAGCCAGCCAGCTGCCCCAGGCCACCTGGGACCAAGACGG 3000
Db 3903 GACACTGGCTGCGCAGGTGTGAGCCAGCCAGCTGCCCCAGGCCACCTGGGACCAAGACGG 3962
QY 3001 A 3001
Db 3963 A 3963

RESULT 10
AX039409
LOCUS AX039409
DEFINITION Sequence 2 from Patent WO0063381.
ACCESSION AX039409
VERSION AX039409.1 GI:11229476
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Zeng W., Stanton L. and Kong H.
TITLE Mammalian protein with putative function in signal transduction
JOURNAL Patent: WO 0063381-A 2 26-OCT-2000;
        SCIOS INC. (US)
FEATURES             Location/Qualifiers
     source            1..4175
                        /organism="Homo sapiens"
                        /mol_type="unassigned DNA"
                        /db_xref="taxon:9606"
                        <1..4056
                        /note="unnamed protein product"
                        /codon_start=1
                        /protein_id="CAC16624.1"
                        /db_xref="GI:11229477"
     translation="GVQFSQYSGSPFVSPFIITQNPVSEASDIWANGVIVSYLSLTGSS
PFAGESDRATLLNLEGRVSWSSPMAHLSEDAKOFIKATLQRAQPAQPSAQCPSLHP
WFLKMPABEAHFINTKQLFLARSRWQSLMSYKILVMRSIPELLRGPDPSPSLG
VARHLCDRTGSSSSSSSSDNELAPFAKSLPPSPVTHSPILLHPRGFRGALAPLPEE
AASRSRTAPAPPEGAPGAPAAQCVPVRSVTRSLFVHQAGSSPRHGAAPGRR
HPARRHLKGGVIAAGLPGLREPLMEHVLFEAAAREFOATLLAKAPSFETALRLPA
SHTHLAPGSHSLHDSPTPRPSSEACGEAQLPSAPSGGAPIDRMGHPQSKQLPS
EAPLVSSPFLGQPAPAPAKAPPLDSKMGFGDISLPGRPKPGPCSSPFSASQASS
SQVSLRVSSQVTEPGPSLDAEGWTQEAEDLSSTPTLQRPQOATMRKFSLGGRG
GYAGVAGYGFAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG
AESPLPQVSARPVPEVGRAPTRSSPEPTPEDIQVSLVQIRDLSDAEADTIISLDI
SEVDPAYNLSDLYDKYLPFEFMI FRKVPKSAQPEPPSPMAEEELAEPEPTWPFPG
ELGPHAGLITESEEDVDKALAAVAARPKWSSPSRSLFHPFGRHLPLDEPAELGLR
ERKVASHEIRILKRGPEKEGPPRRKPGSLASFRSLGKSWDRAPFTFLRELSDER
VYLGQSVTLACQVQAQAATSKDQAPLESSRVLISATLKNLQLLTILVYVABDL
GVYTCVSNALGTVTTLVLRKAERSSPCPDIGEVADGVLLVVKPVSVPVTVI
VOCSLGSSWTTLASDIPDCVLTSTKISRGTYTFRCTACVSKAGMPVSSPSEVLIG
GSHLSEESOGSRSAQPLPTKTFATQIQRGFSVVRQCWEKASGRALAAKIIPIY
HPKDTAVLREYALKGLRHPHQAQHAAYLSFRHLVLIILECSGPELLPGLAERASY
SSEVKYJMWQMLSATQYLNQHILHDLRSNMIIITENYLLKVLVDLNGAQLSDEKV
LPSDFKDYLETMAPELLEGGAVPQTDIWAIGVTAFTIMLSAEYVPSSEGAQLRGL
RKGLVLRRCVAGLSGAVAFRLSTLCAQPMWRPCASSCLQCQPLWTEBGPACSRPAPV
TPTARLRFVRNREKRALLYKRHLAQR"
```

ORIGIN

```
Query Match      88.1%; Score 2643.4; DB 6; Length 4175;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2644; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 357 TGCAGTTTACGCGTACCGCTCCCTGAGTTTCGTTCTCCCGGAGATCATCCAGCAGAAC 416
Db 5 TGCAGTTTACGCGTACCGCTCCCTGAGTTTCGTTCTCCCGGAGATCATCCAGCAGAAC 416
64
```


Db 2225 TGGGCCCCACGCGAGCGCTGGAGATCACAGAGAGGTCAGAGGATGTGGAGCGCGCTGCTGG 2284
Qy 2637 CAGAGGCTGCGTGGGCGAGGAACGCAAGTGTCTCGCGCTCACGCGAGCTCTTCCACT 2696
Db 2285 CAGAGGCTGCGTGGGCGAGGAACGCAAGTGTCTCGCGCTCACGCGAGCTCTTCCACT 2344
Qy 2697 TCCTCTGGGAGGACCTTGGCGCTGGATGAGCTGCGAGCTGGGGCTGCTGAGAGAGTGA 2756
Db 2345 TCCTCTGGGAGGACCTTGGCGCTGGATGAGCTGCGAGCTGGGGCTGCTGAGAGAGTGA 2404
Qy 2757 AGGCTCTCGTGGAGCACATCTCCCGATCTCTGAAGGCGAGGCGGAGGCTTGGAGAGG 2816
Db 2405 AGGCTCTCGTGGAGCACATCTCCCGATCTCTGAAGGCGAGGCGGAGGCTTGGAGAGG 2464
Qy 2817 AGGGGCCCCCAGGAAGAACGAGGCTTGTCTTCCGCTCTCAGGTCTGAAGAGCT 2876
Db 2465 AGGGGCCCCCAGGAAGAACGAGGCTTGTCTTCCGCTCTCAGGTCTGAAGAGCT 2524
Qy 2877 GGGACGAGGCGGACATCTTAAAGGAGCTCTCAGATGAGACTGTGCTGGGCCAGT 2936
Db 2525 GGGACGAGGCGGACATCTTAAAGGAGCTCTCAGATGAGACTGTGCTGGGCCAGT 2584
Qy 2937 CAGTGACACTGGCTGCCAGGTGTGACGCGACGAGTGCAGCCAGGCTGCCAGGCCACCTGGAGCAAAG 2996
Db 2585 CAGTGACACTGGCTGCCAGGTGTGACGCGACGAGTGCAGCCAGGCTGCCAGGCCACCTGGAGCAAAG 2644
Qy 2997 ACGGA 3001
Db 2645 ACGGA 2649

RESULT 11
AB046859
LOCUS AB046859 4041 bp mRNA linear PRI 22-FEB-2001
DEFINITION Homo sapiens mRNA for KIAA1639 protein, partial cds.
ACCESSION AB046859
VERSION AB046859.1 GI:10047354
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
Nagase,T., Kikuno,R., Nakayama,M., Hirose,M., and Ohara,O.
Prediction of the coding sequences of unidentified human genes.
XVIII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro
DNA Res. 7 (4), 273-281 (2000)
20450683
10997877
2 (bases 1 to 4041)
Ohara,O., Nagase,T. and Kikuno,R.
Direct Submission
Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail:cdnaif@kazusa.or.jp).
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-3914)
Location/Qualifiers
1. .4041
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fj06072"
/note="vector:pBluescriptII SK plus"
1. .4041
/gene="KIAA1639"
<1. .3962
/gene="KIAA1639"
/note="Start codon is not identified. fj06072 cDNA clone
for KIAA1639 has a 55-bp deletion at the position between
2845 and 2901, and a 1-bp insertion after the position
1640 of the sequence of KIAA1639, respectively."

/codon_start=3
/product="KIAA1639 protein"
/protein_id="BAB13455.1"
/db_xref="GI:10047355"
/translation="AMGVISYLSLTCSPFAGESDRATLLNVLEGRVSWSPMAAHLR
EDAKDFTKLQAPAPSAACLSHPWFLKMPAEAEHINTKOLFKLIARWR
SLMSYKIPILVMSIPPELLRPPDPSLPSLVARHLCRDTGGSSSSSSSDNDELAPARAK
SLPSPVTHSLPHPRGLRPSLPEAEARESESTAPAPASPEGSGAGAGGCV
RHVSIRSLFVQAGESPEHGALAPGSRHRPARRHLKGGYIAGALFGLREPLMEHRV
LEEAAREBAQTLAKAFSEFATLRLPASGTHLAPGSHSLSHSDSPSTPPSPSACGE
AQLRPSAPSCAPLRDMGHPQKSPSTCGHPTAQPERSPDPMQAPAFCHPQK
GSAPEGCEPHVAVAPCPGSPFGSCKEAPLVSPSPFLGQPOAPAPAKAPLDSK
MGPDILSLRKPXPGPCSPGSAOSSSSSLRVGSSQVTEPGSLDAEGWTSK
EDLSSTPTLQRPQOATMRKFSILGGRGGYAGVAGTFRAGDAGMLGQGPWAI
AWAVOSRESEREAREASQSEARAEPLPOVSARVPVGRVPTRSSPEPTPW
EDTQVSLVQIRLSDGAEAADTSLDISVDPAYLNLSDLYDKYLPFFFMILFRKPY
KWSQPPSPMAEEELAEFPETWPMWPGELGPHAGLEITEESDVEDLALAAVGRK
KSPSRSLSLKSWDRAPTLRLSDTAVLQSVTLACOVSAQPAQAQWSDGAPL
PGLASRLSLGKNDPOLLTLVVVAEDLVYTCVSNALGTVTTTGVLRKAERSSSP
ESRSVLISATLKNFOLLTLVVVAEDLVYTCVSNALGTVTTTGVLRKAERSSSP
CPDGEVYADGVLLVMKPVESYGPVTYVQCSLEGGSWTTLASDIIDPCCYLTSLSRG
GTYPRTACVSKAGMGPSPSPQVLLGSPHSLABEESQGRSAQPLSTFTAFQTQ
IQGRFSVRQCWEKASGRALAAKIIPYHPDKTAVLREYEALGLRHQAHAAY
LSRHLVILLELCSPPELLPCLAEASYSSEVKDYLMQSLQIHLNQHILHLDIR
SENNIITEYLLKVDLGNASQSKYPSDKFYDYLETMELLEGQCAVPQDILWAI
GVTAFTMLSAEYVPSSEGARDLQRLKGLVRLSRCVAGLSGGAVAFRLSTLCAQPMW
RPCASSCLQCPMLTERGPACSRPAPVTFTPARLRFVFRNREKREALLYKHNLAQVR"

ORIGIN

Query Match 85.3%; Score 2561; DB 9; Length 4041;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2561; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 441 GGGCCATGGGTGCATCTCTACCTCAGCTGACCTGCTCATCCCATTTGCCGGGCGAGA 500
Db 1 GGGCCATGGGTGCATCTCTACCTCAGCTGACCTGCTCATCCCATTTGCCGGGCGAGA 60
Qy 501 GTGACCGTGCACCCCTCTGAACTGCTTGGAGGGGGCGGTGTCTATGGAGCAGGCCCATGG 560
Db 61 GTGACCGTGCACCCCTCTGAACTGCTTGGAGGGGGCGGTGTCTATGGAGCAGGCCCATGG 120
Qy 561 CTGCCCCACCTCAGCGAAGAGCGCAAGACTTATCAAGGCTAGCTGCGAGAGAGCCCTTC 620
Db 121 CTGCCCCACCTCAGCGAAGAGCGCAAGACTTATCAAGGCTAGCTGCGAGAGAGCCCTTC 180
Qy 621 AGGCCCGGCTAGTGGCGGCCAGTGCCTCTCCACCCCTGTTCTCGAAATCCATCGCTG 680
Db 181 AGGCCCGGCTAGTGGCGGCCAGTGCCTCTCCACCCCTGTTCTCGAAATCCATCGCTG 240
Qy 681 CGGAGAGGCGCCATTTTCATCAACCAAGCAGCTCAAGTTCTCTTGGGCCGGAAGTCGCT 740
Db 241 CGGAGAGGCGCCATTTTCATCAACCAAGCAGCTCAAGTTCTCTTGGGCCGGAAGTCGCT 300
Qy 741 GGCAGAGTTTCCCTGATGAGTACAAGTCCATCTGCTGATGCGCTCCATCCCTGAGCTGC 800
Db 301 GGCAGAGTTTCCCTGATGAGTACAAGTCCATCTGCTGATGCGCTCCATCCCTGAGCTGC 360
Qy 801 TGGGGGGCCACCCGACAGCCCTCTCCCTGGGGTAGCCCGGACCTCTGCGAGGACACTG 860
Db 361 TGGGGGGCCACCCGACAGCCCTCTCCCTGGGGTAGCCCGGACCTCTGCGAGGACACTG 420
Qy 861 GTGGCTCTCCAGTTCCTCTCTCTTGAACAAGAGCTCGCCCCCATTTGTCGCGGCTA 920
Db 421 GTGGCTCTCCAGTTCCTCTCTCTTGAACAAGAGCTCGCCCCCATTTGTCGCGGCTA 480
Qy 921 AGTCACTGCGCCCTCCCGGTGACACACTCAGCTGTCACCCCGGGGCTTCCTCG 980
Db 481 AGTCACTGCGCCCTCCCGGTGACACACTCAGCTGTCACCCCGGGGCTTCCTCG 540
Qy 981 GGCCTCTCGGCCAGCTGCTGAGGAGGAGCGAGCGCTGAGCGCTCCACCGAGGCCCCAG 1040
Db 541 GGCCTCTCGGCCAGCTGCTGAGGAGGAGCGAGCGCTGAGCGCTCCACCGAGGCCCCAG 600

1041 QY CTCGGCTGCATCTCCGAGGGTCCGGGCGCACCGGCGCCACAGGGCTGCGTGCCTGGCC 1100
601 Db CTCGGCTGCATCTCCGAGGGTCCGGGCGCACCGGCGCCACAGGGCTGCGTGCCTGGCC 660
1101 QY ACAGCGTCAATCCGACGACCTGTTCTACCAACAGGGGGGTGAGAGCCCTCAGACACGGGGCCC 1160
661 Db ACAGCGTCAATCCGACGACCTGTTCTACCAACAGGGGGGTGAGAGCCCTCAGACACGGGGCCC 720
1161 QY TGGCCCCGGGAGCAGGGGGCAGCCCGGCCCGCGCGCGGCGCACTGCTGAAGGGCGGGTACA 1220
721 Db TGGCCCCGGGAGCAGGGGGCAGCCCGGCCCGCGCGCGGCGCACTGCTGAAGGGCGGGTACA 780
1221 QY TTGGGGGGGCGCTGCAGCGCTTCGGCGAGCCACTGATGAGACACCGCGCTGCTGAGGAGG 1280
781 Db TTGGGGGGGCGCTGCAGCGCTTCGGCGAGCCACTGATGAGACACCGCGCTGCTGAGGAGG 840
1281 QY AGGCGCGCAGGGAGGAGCAGGCGCACCTCTCTGGCCAAAGCCCTCATTTGAGAGACTGCC 1340
841 Db AGGCGCGCAGGGAGGAGCAGGCGCACCTCTCTGGCCAAAGCCCTCATTTGAGAGACTGCC 900
1341 QY TCCGGCTGCTGCTCTGGCAACCACTTGGGCCCTTGGCCAGCAGCCACTTCCCTTGAAACATG 1400
901 Db TCCGGCTGCTGCTCTGGCAACCACTTGGGCCCTTGGCCAGCAGCCACTTCCCTTGAAACATG 960
1401 QY ACTCTCGAGACACCCCGCCCTCTCGGAGGCTTCGGTGAAGCAGCAGCTGCTT 1460
961 Db ACTCTCGAGACACCCCGCCCTCTCGGAGGCTTCGGTGAAGCAGCAGCTGCTT 1020
1461 QY CAGCCCCCTCGGGGGGGCCCTTATCAGGGACATGGGGCACCCTCAGGGCTTCCAGAGAGC 1520
1021 Db CAGCCCCCTCGGGGGGGCCCTTATCAGGGACATGGGGCACCCTCAGGGCTTCCAGAGAGC 1080
1521 QY TTCCATCACTGGTGGCCACCCAGGCACTGTCAAGCAGAGAGGCGCATCCCGGACAGCC 1580
1081 Db TTCCATCACTGGTGGCCACCCAGGCACTGTCAAGCAGAGAGGCGCATCCCGGACAGCC 1140
1581 QY CTTGGGGGAGCAGCCCTTCTGCGCACCCAGAGGGTCTGCGCCCGCAGAGGGCT 1640
1141 Db CTTGGGGGAGCAGCCCTTCTGCGCACCCAGAGGGTCTGCGCCCGCAGAGGGCT 1200
1641 QY GCAGCCCCCAGCAGCAGTGGCCCATGCCCCCTCTGGCTCTCTCCCTCCAGGATCTTGCA 1700
1201 Db GCAGCCCCCAGCAGCAGTGGCCCATGCCCCCTCTGGCTCTCTCCCTCCAGGATCTTGCA 1260
1701 QY AAGAGGCCCCCTTAGTACCTCAAGCCCCCTTTTGGGACAGCCCCAGGACCCCTTGCCC 1760
1261 Db AAGAGGCCCCCTTAGTACCTCAAGCCCCCTTTTGGGACAGCCCCAGGACCCCTTGCCC 1320
1761 QY GTGCCAAGCAAGCCCCCATTTGACATCTAAGATGGGGGCTTGAGACATCTCTTCTCTG 1820
1321 Db GTGCCAAGCAAGCCCCCATTTGACATCTAAGATGGGGGCTTGAGACATCTCTTCTCTG 1380
1821 QY GGAGGCCAAAACCGGGCCCTTGAGTTCGCCAGGGTCAAGCTTCCAGGCGAGCTCTTCCC 1880
1381 Db GGAGGCCAAAACCGGGCCCTTGAGTTCGCCAGGGTCAAGCTTCCAGGCGAGCTCTTCCC 1440
1881 QY AAGTGAAGCTCCCTCAGGGTGGGCTCTCTCCAGGTGGGACAGAGCTTGGCCCCCTTCTGG 1940
1441 Db AAGTGAAGCTCCCTCAGGGTGGGCTCTCTCCAGGTGGGACAGAGCTTGGCCCCCTTCTGG 1500
1941 QY ATGGGAGGGCTGAGCCAGAGGCTGAGGATCTGTCCGACTCCACACCCACCTTGGAGC 2000
1501 Db ATGGGAGGGCTGAGCCAGAGGCTGAGGATCTGTCCGACTCCACACCCACCTTGGAGC 1560
2001 QY GGCCTCAGGAACAGCGACCATGCGCAAGTTCTCCCTGGGTGGTCCGGGGGCTACGAG 2060
1561 Db GGCCTCAGGAACAGCGACCATGCGCAAGTTCTCCCTGGGTGGTCCGGGGGCTACGAG 1620
2061 QY CGCTGGCTGGCTATGGCACTTGTGCTTTGGTGGAGATGCAAGGGGCGATGCTGGGGCAGG 2120
1621 Db CGCTGGCTGGCTATGGCACTTGTGCTTTGGTGGAGATGCAAGGGGCGATGCTGGGGCAGG 1680
2121 QY GGCCCATGTGGGCCAGGATAGCTGGGCTGTGTCCAGCTCGGAGGAGGAGGAGGAGG 2180

1581 Db GGCCCATGTGGGCCAGGATAGCTTGGCTGTGTCCAGTCCGAGGAGGAGGAGGAGGAGG 1740
2181 QY AGGCCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGCGCAGGGCTGAGAGCCCACTGC 2240
1741 Db AGGCCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGCGCAGGGCTGAGAGCCCACTGC 1800
2241 QY CCAGAGTCAAGTCAAGGGCTGTGCTGAGTCCGAGGAGGCTCCACACAGGAGCTCTCCAG 2300
1801 Db CCAGAGTCAAGTCAAGGGCTGTGCTGAGTCCGAGGAGGCTCCACACAGGAGCTCTCCAG 1860
2301 QY AGCCACCCCATGGAGGACATCCGGCAGGTCTCCCTGGTGCAGATCCGGGACCTGTGAG 2360
1861 Db AGCCACCCCATGGAGGACATCCGGCAGGTCTCCCTGGTGCAGATCCGGGACCTGTGAG 1920
2361 QY GTGATGCGAGGCGGCGGACAAATATCCCTGGGCAATTTCCGAGGTGAGACCCCGCTTACC 2420
1921 Db GTGATGCGAGGCGGCGGACAAATATCCCTGGGCAATTTCCGAGGTGAGACCCCGCTTACC 1980
2421 QY TCAACCTCTCAGACTGTACGATATCAAGTACCTCCCATTTGATGATCTTCAGGA 2480
1981 Db TCAACCTCTCAGACTGTACGATATCAAGTACCTCCCATTTGATGATCTTCAGGA 2040
2481 QY AAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCATGCTGAGGAGGAGCTGGCGG 2540
2041 Db AAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCATGCTGAGGAGGAGCTGGCGG 2100
2541 QY AGTTCCCGGAGCCCACTGGGCCCTGGCCAGGTGAATCTGGGGCCCCCAGCAGGCTTGGAGA 2600
2101 Db AGTTCCCGGAGCCCACTGGGCCCTGGCCAGGTGAATCTGGGGCCCCCAGCAGGCTTGGAGA 2160
2601 QY TCACAGAGGAGTCAAGAGATGTGACCGGCTGTGCGCAGAGGCTGCGTGGGCGAGGAGC 2660
2161 Db TCACAGAGGAGTCAAGAGATGTGACCGGCTGTGCGCAGAGGCTGCGTGGGCGAGGAGC 2220
2661 QY GCAAGTGTCTCCGCTCAGCAGCAGCTTCTTCCACTTCCCTGGGAGGAGCAGCTGCGCGCTGG 2720
2221 Db GCAAGTGTCTCCGCTCAGCAGCAGCTTCTTCCACTTCCCTGGGAGGAGCAGCTGCGCGCTGG 2280
2721 QY ATGAGCTTGCAGAGCTGGGGCTGCGTGAAGAGTGAAGGCTTCCGTGGAGCAGCATCTCCC 2780
2281 Db ATGAGCTTGCAGAGCTGGGGCTGCGTGAAGAGTGAAGGCTTCCGTGGAGCAGCATCTCCC 2340
2781 QY GGATCTTGAAGGAGGCGGAGAGGCTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGCAG 2840
2341 Db GGATCTTGAAGGAGGCGGAGAGGCTTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGCAG 2400
2841 QY GCCTTGTCTTCCCTCCGCTCTCAGGTCTGAGAGAGTGGGACCGGAGCGCCGACATTCCTAA 2900
2401 Db GCCTTGTCTTCCCTCCGCTCTCAGGTCTGAGAGAGTGGGACCGGAGCGCCGACATTCCTAA 2460
2901 QY GGGAGCTTCTCAGATGAGATGTGTGCTTCCCTGGGCCAGTCAAGTGAACATGCGCTGCCAGGTGT 2960
2461 Db GGGAGCTTCTCAGATGAGATGTGTGCTTCCCTGGGCCAGTCAAGTGAACATGCGCTGCCAGGTGT 2520
2961 QY CAGCCCAAGCAGCTGCCCCAGGCCACCTGGAGCAAGACGGA 3001
2521 Db CAGCCCAAGCAGCTGCCCCAGGCCACCTGGAGCAAGACGGA 2561

RESULT: 12
AL670729
LOCUS Human DNA sequence from clone RP11-245P10 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL670729
VERSION AL670729.19 GI:22316158
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 62164)

AUTHORS
TITLE
JOURNAL

Van Hellmond, Z.
Direct Submission
Submitted (15-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Aug 19, 2002 this sequence version replaced gi:21727388.

COMMENT

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclones; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/Chr1>
RP11-245P10 is from the library RPci-11.1 constructed by the group
of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>

FEATURES

source
Location/Qualifiers
1..62164
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-245P10"
/clone_lib="RPci-11.1"

ORIGIN

Query Match 69.0%; Score 2070.4; DB 9; Length 62164;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2077; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 719 TTCTCTCTGGCCCGAAGTCGCTGGCAGCGTTCCCTGATGAGCTACAAAGTCATCCTCGTG 778
Db 24359 TGCCTGTGCCCCCATCGCGTTCCAGCGTTCCCTGATGAGTACAAAGTCCATCCTGGTG 24418

Qy 779 ATGGCTCCATCCCTGAGCTGCTCGGGGGCCACCCGACAGCCCTCCCTCGGCGTAGCC 838
Db 24419 ATGGCTCCATCCCTGAGCTGCTCGGGGGCCACCCGACAGCCCTCCCTCGGCGTAGCC 24478

Qy 839 CGGCACCTCTGACGGACACTGGTGGCTCTCCAGTTCTCTCCCTCTGACACAGCAG 898
Db 24479 CGGCACCTCTGACGGACACTGGTGGCTCTCTCCAGTTCTCTCCCTCTGACACAGCAG 24538

Qy 899 CTGCCCCCATTTGCCCGGGCTAAGTCACTGCACACCTCCCGGGTGACACACTCACCACCTG 958
Db 24539 CTGCCCCCATTTGCCCGGGCTAAGTCACTGCACACCTCCCGGGTGACACACTCACCACCTG 24598

Qy 959 CTGCACCCCGGGGCTTCTCGGCCCTCGGCCAGCCTGCTGAGGAAGCCGAGGCCAGT 1018
Db 24599 CTGCACCCCGGGGCTTCTCGGCCCTCGGCCAGCCTGCTGAGGAAGCCGAGGCCAGT 24658

Qy 1019 GAGCGCTCCACCGAGGCCACCGAGCTCCCGCTGCATCTCCCGAGGGTGGCCGGCCACCGGCC 1078

Db 24659 GAGCGCTCCACCGAGGGCCCCAGCTCCGCTGCATCTCCCGAGGGTGGCCGGSCC 24718
Qy 1079 GCGCAGGGCTGGCTGGCCCCCGGCACAGCGTCACTCCGAGCCTGTCTTACCACAGCGCGGT 1138
Db 24719 GCGCAGGGCTGGCTGGCCCCCGGCACAGCGTCACTCCGAGCCTGTCTTACCACAGCGCGGT 24778
Qy 1139 GAGAGCCCTGAGCACGGGGGCGCTGGCCCCCGGGGAGCAGGGGACACCGCGCCGCGCGG 1198
Db 24779 GAGAGCCCTGAGCACGGGGGCGCTGGCCCCCGGGGAGCAGGGGACACCGCGCCGCGCGG 24838
Qy 1199 CACCTGCTGAAGGGCGGTACATTTGCGGGGCGCTGCCAGCGCCTTGCAGGCCACTGATG 1258
Db 24839 CACCTGCTGAAGGGCGGTACATTTGCGGGGCGCTGCCAGCGCCTTGCAGGCCACTGATG 24898
Qy 1259 GAGCACCGCTGCTGGAGGAGGGCGCCGAGGGAGAGCAGGCCACCTCTCTGGCCAAA 1318
Db 24899 GAGCACCGCTGCTGGAGGAGGGCGCCGAGGGAGAGCAGGCCACCTCTCTGGCCAAA 24958
Qy 1319 GCCCCTCATTTGAGAGCTGCCCTCCGGCTGCTGCCCTTCTGGCACCCCACTTTGGCCCTGGC 1378
Db 24959 GCCCCTCATTTGAGAGCTGCCCTCCGGCTGCTGCCCTTCTGGCACCCCACTTTGGCCCTGGC 25018
Qy 1379 CACAGCCACTCTTGGAAACATGACTCTTCCGAGCACCCCCCGCTCTCTCGAGGCGCTGC 1438
Db 25019 CACAGCCACTCTTGGAAACATGACTCTTCCGAGCACCCCCCGCTCTCTCGAGGCGCTGC 25078
Qy 1439 GGTGAGGCACAGGAGCTGCTTCCAGCCCTTCCGGGGGGGGCCCTATCAGGGACATGGGG 1498
Db 25079 GGTGAGGCACAGGAGCTGCTTCCAGCCCTTCCGGGGGGGGCCCTATCAGGGACATGGGG 25138
Qy 1499 CACCTCTCAGGGCTCCAAGCAGCTTCCATCCACTGTGGTGGCCACCCAGGCACTGTCTCAGCCA 1558
Db 25139 CACCTCTCAGGGCTCCAAGCAGCTTCCATCCACTGTGGTGGCCACCCAGGCACTGTCTCAGCCA 25198
Qy 1559 GAGAGGCACTCCCGGACAGCCTTGGGGGACGACAGCCCTTCTTGCACACCCCAAGCAG 1618
Db 25199 GAGAGGCACTCCCGGACAGCCTTGGGGGACGACAGCCCTTCTTGCACACCCCAAGCAG 25258
Qy 1619 GGTTCCTGCCCCCAGAGAGGGCTGCAGCCCCCAGCCAGAGTTCGCCCTTCCCTCTGGC 1678
Db 25259 GGTTCCTGCCCCCAGAGAGGGCTGCAGCCCCCAGCCAGAGTTCGCCCTTCCCTCTGGC 25318
Qy 1679 TCCTTCCCTCCAGGATCTTGCAGAGGGCCCCCTTAGTACCTTCAAGCCCTTCTTGGGA 1738
Db 25319 TCCTTCCCTCCAGGATCTTGCAGAGGGCCCCCTTAGTACCTTCAAGCCCTTCTTGGGA 25378
Qy 1739 CAGCCCCAGGACCCCTTGGCCCTTGCAGAGCCGCCCATTTGAGCTCTTAAGATGGGG 1798
Db 25379 CAGCCCCAGGACCCCTTGGCCCTTGCAGAGCCGCCCATTTGAGCTCTTAAGATGGGG 25438
Qy 1799 CTTGGAGACATCTCTCTCTGGAGGCCCAAAACCCGGCCCTTGCAGTTCCTCCAGGGTCA 1858
Db 25439 CTTGGAGACATCTCTCTCTGGAGGCCCAAAACCCGGCCCTTGCAGTTCCTCCAGGGTCA 25498
Qy 1859 GCTTCCAGCGAGCTCTTCCCAAGTGAAGTCCCTCAGGGTGGGCTCTCTCCAGGTTGGG 1918
Db 25499 GCTTCCAGCGAGCTCTTCCCAAGTGAAGTCCCTCAGGGTGGGCTCTCTCCAGGTTGGG 25558
Qy 1919 ACAGAGCTTGGCCCTTCCCTGATGCGAGGGCTGAGCCAGAGGCTGAGGATCTGTCC 1978
Db 25559 ACAGAGCTTGGCCCTTCCCTGATGCGAGGGCTGAGCCAGAGGCTGAGGATCTGTCC 25618
Qy 1979 GACTCCACACCACTTTCAGCGGGCTCAGGAAACAGGGCCACATGCGCAAGTTCTCCCTG 2038
Db 25619 GACTCCACACCACTTTCAGCGGGCTCAGGAAACAGGGCCACATGCGCAAGTTCTCCCTG 25678
Qy 2039 GGTGTGCGGGGGCTACGACAGCGTGGCTGTGCTATGGCACTTTGCCCTTTGGTGGAGAT 2098
Db 25679 GGTGTGCGGGGGCTACGACAGCGTGGCTGTGCTATGGCACTTTGCCCTTTGGTGGAGAT 25738
Qy 2099 GCAGGGGCTGCTGGGGCAGGGGCCCTATGTGGGCCAGGATAGCTGGGCTGTGTCCAG 2158

Db 25739 GCAGGGGCGATGCTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGTCCTCCAG 25798
Qy 2159 TCGAGGAGGAGGAGCAGAGAGGCCAGGGCTGATGCTCCAGTTCGAGGAGCAGCAGAG 2218
Db 25799 TCAGAGGAGGAGGAGCAGAGAGGCCAGGGCTGATGCCAGTCCAGTCCGAGGAGCAGCAGAG 25858
Qy 2219 GCCAGGGCTGAGAGGCCACTGCCAGGTCTAGTCAAGGCTGTGCCCTGAGGTCGGCAGG 2278
Db 25859 GCCAGGGCTGAGAGGCCACTGCCAGGTCTAGTCAAGGCTGTGCCCTGAGGTCGGCAGG 25918
Qy 2279 GCTCCACAGGAGCTCTCCAGAGGCCACCCATGGGAGGACATCGGGCAGGTCTCCCTG 2338
Db 25919 GCTCCACAGGAGCTCTCCAGAGGCCACCCATGGGAGGACATCGGGCAGGTCTCCCTG 25978
Qy 2339 GTGAGATCCGGGAGCTCTGAGGTGATGCGAGAGCGCGCAGCAATATCTCCCTGACATTT 2398
Db 25979 GTGAGATCCGGGAGCTCTGAGGTGATGCGAGAGCGCGCAGCAATATCTCCCTGACATTT 26038
Qy 2399 TCCGAGGTGGAGCCCGCTACTCAACCTCTCAGACCTGACGATATCAAGTACCTCCCA 2458
Db 26039 TCCGAGGTGGAGCCCGCTACTCAACCTCTCAGACCTGACGATATCAAGTACCTCCCA 26098
Qy 2459 TTCGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCCGCCCTCCCCC 2518
Db 26099 TTCGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCCGCCCTCCCCC 26158
Qy 2519 ATGCTGAGGAGGAGCTGGCGAGTTTCCGAGAGCCACGTGGCCCTGCGCCAGGTGAATG 2578
Db 26159 ATGCTGAGGAGGAGCTGGCGAGTTTCCGAGAGCCACGTGGCCCTGCGCCAGGTGAATG 26218
Qy 2579 GGGCCCCACGAGCGCTGGAGATCACAGAGAGTCAGAGGATGTGGAGCGCGCTGTGGCA 2638
Db 26219 GGGCCCCACGAGCGCTGGAGATCACAGAGAGTCAGAGGATGTGGAGCGCGCTGTGGCA 26278
Qy 2639 GAGGCTGCGTGGGAGGAGCGAAGTGGTCTCGCGCTGCAAGCAGGCTCTTCCACTTC 2698
Db 26279 GAGGCTGCGTGGGAGGAGCGAAGTGGTCTCGCGCTGCAAGCAGGCTCTTCCACTTC 26338
Qy 2699 CCTGGAGGACCTGCGCGCTGGATGAGCCTGAGAGTGGGGCTGCTGAGAGAGTGAAG 2758
Db 26339 CCTGGAGGACCTGCGCGCTGGATGAGCCTGAGAGTGGGGCTGCTGAGAGAGTGAAG 26398
Qy 2759 GCCTCCGTGGAGACATCTCCGGATCTGAGGCGAGGCGGAGGT 2806
Db 26399 GCCTCCGTGGAGACATCTCCGGATCTGAGGCGAGGCGGAGGT 26446

RESULT 13
LOCUS CQ730656 18524 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 16590 from Patent W002068579.
ACCESSION CQ730656
VERSION CQ730656.1 GI:42305092
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kirs, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 16590 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
Location/Qualifiers
1..18524
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 57.9%; Score 1738.8; DB 6; Length 18524;

Best Local Similarity 82.2%; Pred. No. 8.5e-270;
Matches 2277; Conservative 0; Mismatches 37; Indels 456; Gaps 6;
Qy 257 CTGACATAAAGCCCTCTAAACATCTGATGTGATCTCTGCGGGAAGACATTTAAATC 316
Db 14805 CAGCCCTTTCAGCCCTCTTAACATCTGATGTGATCTCTGCGGGAAGACATTTAAATC 14864
Qy 317 TGCACATTTGGCTTTGCCAGAACATCACCCAGCAGAGCTGCAGTTTCAGCCAGTACGGC 376
Db 14865 TGCACATTTGGCTTTGCCAGAACATCACCCAGCAGAGCTGCAGTTTCAGCCAGTACGGC 14924
Qy 377 TCCCTGATGTTGCTCTCCCGAGATCATCAGCAGAACCTGTGAGAGGAGCCTCCGAC 436
Db 14925 TCCCTGATGTTGCTCTCCCGAGATCATCAGCAGAACCTGTGAGAGGAGCCTCCGAC 14984
Qy 437 ATTTGGGCCATGGGTGATCTCTTACCTCAGCCTGACCTGCTCATCCCCATTTGCCGGC 496
Db 14985 ATTT-----GCCTGACCTCTCATCCCCATTTGCCGGC 15017
Qy 497 GAGAGTACCGTGGCACTCTCTGAAAGTCTCTGAGGGGCGGTGTCTATGAGCAGCCCC 556
Db 15018 GAGAGTACCGTGGCACTCTCTGAAAGTCTCTGAGGGGCGGTGTCTATGAGCAGCCCC 15077
Qy 557 ATGCTGCCACCTCAGCAGAGAGCCAAAGCTTTCATCAAGGCTACGCTGCAGAGGCC 616
Db 15078 ATGCTGCCACCTCAGCAGAGAGCCAAAGCTTTCATCAAGGCTACGCTGCAGAGGCC 15137
Qy 617 CCTCAGG----CCGGGCTAGTGGGGCCAGTGGCTCTTCCCACCCCTGG----- 661
Db 15138 CCTCAGGAAATCCAGGGCTGGGGCTGCCAGAGGGCTTCCATCATGTGGAGGGGGCCCTA 15197
Qy 662 -----TTCTGAAATCCATG 676
Db 15198 TCAGGGCACCATGGTCTCTCCATGCTGGCTGGCCACCTCGGTCTCCAGAAATCCATG 15257
Qy 677 CTTGGAGGAGGAGCCACTTTCATCAACAGCAGAGCTCAAGTTCCTCTGCGCCGAAAGT 736
Db 15258 CTTGGAGGAGGAGCCACTTTCATCAACAGCAGAGCTCAAGTTCCTCTGCGCCGAAAGT 15317
Qy 737 CGCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTCTGGTGTGCTGCTCATCCCTGAG 796
Db 15318 CGCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTCTGGTGTGCTGCTCATCCCTGAG 15377
Qy 797 CTGCTGGGGGCCACCGGACAGCCCTCTCTCGGCTAGCCCGGCACTCTGAGAGGAC 856
Db 15378 CTGCTGGGGGCCACCGGACAGCCCTCTCTCGGCTAGCCCGGCACTCTGAGAGGAC 15437
Qy 857 ACTGGTGGCTCTCCAGTTCTCTCTCTGACAGAGCTCGCCCCCATTTTGGCCGG 916
Db 15438 ACTGGTGGCTCTCCAGTTCTCTCTCTGACAGAGCTCGCCCCCATTTTGGCCGG 15497
Qy 917 GCTAAGTCACTGGCACACCTCTCCCGGTGACACACTCACCACCTGCTGCACCCCGGGGCTTC 976
Db 15498 GCTAAGTCACTGGCACACCTCTCCCGGTGACACACTCACCACCTGCTGCACCCCGGGGCTTC 15557
Qy 977 CTGGGGCCTCGGCGAGCTGCTGAGAGAGCCGAGGCACTGAGCGCTCCACCGAGGCC 1036
Db 15558 CTGGGGCCTCGGCGAGCTGCTGAGAGAGCCGAGGCACTGAGCGCTCCACCGAGGCC 15617
Qy 1037 CCAGCTCGCTGCTGCTCTCCGAGGCTGCGGGCCACCGGGCCCGGAGGCTGGTGGCC 1096
Db 15618 CCAGCTCGCTGCTGCTCTCCGAGGCTGCGGGCCACCGGGCCCGGAGGCTGGTGGCC 15677
Qy 1097 CGGCAAGCGCTCATCCCGAGCCTTCTTACCAACAGGCGGTGAGAGCCCTGAGCACGGG 1156
Db 15678 CGGCAAGCGCTCATCCCGAGCCTTCTTACCAACAGGCGGTGAGAGCCCTGAGCACGGG 15737
Qy 1157 GCTTGGCCCGGGGAGAGGCGGCACTCGGCGCGGGCGGCACTTGTGTAAGGGCGGG 1216
Db 15738 GCTTGGCCCGGGGAGAGGCGGCACTCGGCGCGGGCGGCACTTGTGTAAGGGCGGG 15797
Qy 1217 TACATTTGGGGGGGCTGCGCAGGCTGCGCAGGCACTGATGGAGCACCGGCTGCTGAG 1276

Db 15798 TACATTGCGGGGGCTGCCAGGCTGCGCGAGCCACTGATGAGACACCGCGTGTGGAG 15857
Qy 1277 GAGAGGCCCCAGGAGGAGGAGGAGCCACCTCTCTGGCCAAAGCCCCCTCATTCGAGACT 1336
Db 15858 GAGAGGCCCCAGGAGGAGGAGGAGCCACCTCTCTGGCCAAAGCCCCCTCATTCGAGACT 15917
Qy 1337 GCCTCTCGGCTGCTGCTCTGCGACCCACTTGGCCCTTGGCCCAAGCCACTCCCTGGAA 1396
Db 15918 GCCTCTCGGCTGCTGCTCTGCGACCCACTTGGCCCTTGGCCCAAGCCACTCCCTGGAA 15977
Qy 1397 CATGACTCTCCGAGCACCCTCCCTCTCTGGAGGCTCTGGGTGAGGCAAGCGACTG 1456
Db 15978 CATGACTCTCCGAGCACCCTCCCTCTCTGGAGGCTCTGGGTGAGGCAAGCGACTG 16019
Qy 1457 CTTTCAGCCCTCTCCGAGGAGGCTCTCTATCAGGCAATGGGGCACTCCAGGGCTCCAAG 1516
Db 16020 ----- 16019
Qy 1517 CAGCTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGAGAGGCAATCCCGGAC 1576
Db 16020 ----- 16019
Qy 1577 AGCCCTTGGGGGAGCCAGCCCTTTCTGCCACCCCAAGCAGGGTTCTGCCCCAGGAG 1636
Db 16020 ----- 16019
Qy 1637 GCGTGCAGCCCCCAGCCAGCACTTGGCCCAATGCCCTCTCTGGCTCTTCCCTCCAGGATCT 1696
Db 16020 -----GGATCT 16025
Qy 1697 TGCAAGAGGCCCTTAGTACCTCAAGCCCTCTTGGGACAGCCCCAGGCAAGCCCT 1756
Db 16026 TGCAAGAGGCCCTTAGTACCTCAAGCCCTCTTGGGACAGCCCCAGGCAAGCCCT 16085
Qy 1757 GCGCTTGCACAAAGCAAGCCCTTGGACTCTTAAGATGGGCTCTGGAGACATCTCTTT 1816
Db 16086 GCGCTTGCACAAAGCAAGCCCTTGGACTCTTAAGATGGGCTCTGGAGACATCTCTTT 16145
Qy 1817 CTTGGAGGCCCAAAACCGGCCCCCTGAGTTCCCGCAGGGTCAGCTCCAGGCGAGCTCT 1876
Db 16146 CTTGGAGGCCCAAAACCGGCCCCCTGAGTTCCCGCAGGGTCAGCTCCAGGCGAGCTCT 16193
Qy 1877 TCCCAAGTGAGCTCCCTCAGGGTGGGCTCTCCAGGTCGCGACAGAGCCCTGCCCTCC 1936
Db 16194 -----GAGCTTGGCCCCCTCC 16208
Qy 1937 CTGATGCGGAGGGCTGGAACCAAGAGCTGAGGATCTGTCGACTCCACACCCACCTTG 1996
Db 16209 CTGATGCGGAGGGCTGGAACCAAGAGCTGAGGATCTGTCGACTCCACACCCACCTTG 16268
Qy 1997 CAGCGGCTCAGGAACAGGCGACCAATGGCAAGTTCTCCCTGGTGTGCGGGGGCTAC 2056
Db 16269 CAGCGGCTCAGGAACAGGCGACCAATGGCAAGTTCTCCCTGGTGTGCGGGGGCTAC 16328
Qy 2057 GCAGCGTGGCTGATGCACTTTGCCCTTTGGTGGAGATGAGGGGGCATGCTGGGG 2116
Db 16329 GCAGCGTGGCTGATGCACTTTGGCTTTGGTGGAGATGAGGGGGCATGCTGGGG 16388
Qy 2117 CAGGGCCCATGTTGGCCAGGATGAGCTGGCTGTGTCAGTTCGAGGAGGAGGAGCAG 2176
Db 16389 CAGGGCCCATGTTGGCCAGGATGAGCTGGCTGTGTCAGTTCGAGGAGGAGGAGCAG 16448
Qy 2177 GAGGAGCCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGCCAGGGCTGAGAGCCCA 2236
Db 16449 GAGGAGCCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGCCAGGGCTGAGAGCCCA 16508
Qy 2237 CTGCCCCAGGTCTAGTCAAGGCTGTGCTGAGGTGCGGAGGCTCCACAGGAGCTCT 2296
Db 16509 CTGCCCCAGGTCTAGTCAAGGCTGTGCTGAGGTGCGGAGGCTCCACAGGAGCTCT 16568
Qy 2297 CCAGAGCCACCCATGGGAGGACATCGGAGGCTCTCCCTGTGTGAGATTCGGGACCTG 2356
Db 16569 CCAGAGCCACCCATGGGAGGACATCGGAGGCTCTCCCTGTGTGAGATTCGGGACCTG 16628

Qy 2357 TCAGTGATGCGGAGGCGCGCCGACACATATCCCTTGACATTTCCGAGGTGGACCCCGCC 2416
Db 16629 TCAGTGATGCGGAGGCGCGCCGACACATATCCCTTGACATTTCCGAGGTGGACCCCGCC 16688
Qy 2417 TACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTCGAGTTTATCATCTTC 2476
Db 16689 TACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTCGAGTTTATCATCTTC 16748
Qy 2477 AGGAAAGTCCCAAGTCCGCTAGCCAGAGCCGCTTCCCTCATTCGCTGAGGAGGAGCTG 2536
Db 16749 AGGAAAGTCCCAAGTCCGCTAGCCAGAGCCGCTTCCCTCATTCGCTGAGGAGGAGCTG 16808
Qy 2537 GCCGAGTTCGCGAGCCAGCTGGCCCTGGCCAGGTGAATCTGGCCCCCAGCAGGCTG 2596
Db 16809 GCCGAGTTCGCGAGCCAGCTGGCCCTGGCCAGGTGAATCTGGCCCCCAGCAGGCTG 16868
Qy 2597 GAGATCAGAGGAGTCAAGGAGTGGACGCGCTGCTGGCAGAGGCTGCGTGGGCGAG 2656
Db 16869 GAGATCAGAGGAGTCAAGGAGTGGACGCGCTGCTGGCAGAGGCTGCGTGGGCGAG 16928
Qy 2657 AAGCGCAAGTGTCTTCCGCTCAGGAGCTTTCACCTTCCCTGGGAGGCACTTGCCTG 2716
Db 16929 AAGCGCAAGTGTCTTCCGCTCAGGAGCTTTCACCTTCCCTGGGAGGCACTTGCCTG 16988
Qy 2717 CTGATGAGCTCAGAGCTGGGCTGCGTGGAGAGTGAAGGCTCCCTGGAGCAGATC 2776
Db 16989 CTGATGAGCTCAGAGCTGGGCTGCGTGGAGAGTGAAGGCTCCCTGGAGCAGATC 17016
Qy 2777 TCCCGGATCTGAAGGCGAGCGGAAAGTCTGGAGAGGAGGGGCCCCCAGGAAGAAG 2836
Db 17017 -----GTCTGGAGAGGAGGGGCCCCCAGGAAGAAG 17048
Qy 2837 CAGGCTTGTCTTCCGCTCTCAGCTTGAAGAGCTGGAGCCGAGGCGGACATTC 2896
Db 17049 CAGGCTTGTCTTCCGCTCTCAGCTTGAAGAGCTGGAGCCGAGGCGGACATTC 17108
Qy 2897 CTAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCGAGTCAAGTACACTGCGCTGCCAG 2956
Db 17109 CTAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCGAGTCAAGTACACTGCGCTGCCAG 17168
Qy 2957 GTGTACGCCC 2966
Db 17169 ACGAGGCCCC 17178

RESULT 14
AX642969
LOCUS AX642969 3225 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 46 from Patent WO01096547.
ACCESSION AX642969
VERSION AX642969.1 GI:28550118
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Yue, H., Lal, P., Bandman, O., Borowsky, M., Au-Young, J., Lu, Y.,
Gandhi, A.R., Tribouley, C.M., Walla, N., Yao, M.G., Lu, D.A.,
Greenwald, S.R., Ramkumar, J., Griffin, J.A., Kearney, L., Burford, N.,
Nguyen, D.B., Tang, Y.T., Baughn, M.R., He, A., Thornton, M.,
Hafalia, A., Patterson, C., Gururajan, R., Lo, T.P., Khan, F.,
Recipon, S.A., Azimzai, Y., Policky, J.L., Ding, L., Grether, M.,
Elliott, V.S., Thangavelu, K., Batra, S. and Ison, C.H.
HUMAN KINASES
TITLE Patent: WO 01096547-A 46 20-DEC-2001;
JOURNAL Incyte Genomics, Inc. (US)
FEATURES Location/Qualifiers
source 1..3225
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

COMMENT

On Jul 7, 2000 this sequence version replaced qi:8748947.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC

Web site: <http://genome.wustl.edu/qsc/index.shtml>

----- Project Information -----

Center project name: H_NH0661B12

----- Summary Statistics -----
 Remarking vector: W12: 100%

Sequencing vector: M13; 100%
Sequencing vector: pIasmid; 0%

Chemistry: Dye-primer ET; 100%

Chemistry: Dye-terminator Big Dye; 0% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 163945 bases at least

Consensus quality: 167601 bases at least
Consensus quality: 169687 bases at least

Insert size: 148000; agarose-fp

Insert size: 173412; sum-of-contigs

Quality coverage: 5.67 in Q20 ba
Quality coverage: 5.55 in Q20 ba

Quality coverage: 3:33 in Q20 bases -----

* NOTE: This is a 'working draft' sequence. It curr

* consists of 13 contigs. The true order of the pie

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

*	1	2478:	contig	of 2478 bp in length
*	2479	2578:	gap of	unknown length
*	2579	6948:	contig of	4370 bp in length
*	6949	7048:	gap of	unknown length
*	7049	9373:	contig of	2325 bp in length
*	9374	9473:	gap of	unknown length
*	9474	13303:	contig of	3830 bp in length
*	13304	13403:	gap of	unknown length
*	13404	18270:	contig of	4867 bp in length
*	18271	18370:	gap of	unknown length
*	18371	25589:	contig of	7219 bp in length
*	25590	25689:	gap of	unknown length
*	25690	34815:	contig of	9126 bp in length
*	34816	34915:	gap of	unknown length
*	34916	45742:	contig of	10827 bp in length
*	45743	45842:	gap of	unknown length
*	45843	62987:	contig of	17145 bp in length
*	62988	63087:	gap of	unknown length
*	63088	87584:	contig of	24497 bp in length
*	87585	87684:	gap of	unknown length
*	87685	114765:	contig of	27081 bp in length
*	114766	114865:	gap of	unknown length
*	114866	142334:	contig of	27469 bp in length
*	142335	142434:	gap of	unknown length
*	142435	174612:	contig of	32178 bp in length

FEATURES

```
Location/Qualifiers
1. .174612
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="1"
   /clone="RP11-661B12"
```

misc_feature	1. .2478	/note="assembly_name:Contig4"
misc_feature	2579. .6948	/note="assembly_name:Contig5"
misc_feature	7049. .9373	/note="assembly_name:Contig6"
misc_feature	9474. .13303	/note="assembly_name:Contig7"
misc_feature	13404. .18270	/note="assembly_name:Contig8"
misc_feature	18371. .25589	

			/note="assembly_name:Contig9" 25690..34815 /note="assembly_name:Contig10" 34916..45742 /note="assembly_name:Contig11" 45843..62987 /note="assembly_name:Contig12 clone_end:SP6 vector_side:right" 63088..87584 /note="assembly_name:Contig13" 87685..114765 /note="assembly_name:Contig14" 114866..142334 /note="assembly_name:Contig15" 142435..174612 /note="assembly_name:Contig16"		
misc_feature					
misc_feature					
misc_feature					
misc_feature					
misc_feature					
misc_feature					
ORIGIN					
Query Match	50.5%;	Score 1517;	DB 2;	Length 174612;	
Best Local Similarity	97.2%;	Pred. No. 2e-234;			
Matches 1641;	Conservative	0;	Mismatches 35;	Indels 13;	Gaps 9;
QY	1118	CTGTTTCACCA	CAGCGGGTGAGAGCCCTGTAGACA	CGGGGCCCTTGGCCCCCGGGAGCAGG	1177
Db	25403	CTTTTCTACC	ACCAAGCGTGATGAACCCCT--	-GCAGGGCGCTCTGGCTCCGGGCAACAAG	25346
QY	1178	CGGCACCCGCG	CGGGCGGCACCTGTGAAGGCGGGTACATT	TGGGGGGCGCTGCACA	1237
Db	25345	CGAAACCCGG-	GCACAGTCGGCACCTGTCTGAAGGGCGGGTACATT	TGCGGG--GCTTGCCA	25289
QY	1238	GGCCTCGCGAG	CCACTGATGAGACACCGCTGTCTGAGAGAGGAGCG	CGCCAGGAGGAG	1297
Db	25288	GGCCTCGCGAG	CCACTGATGAGAGNACACGCTGCTGAGAGAGGAGCG	CGCCAGGAGGAG	25229
QY	1298	CAGGCCACCTCT	CTGTGCCAAAAGCCCCTCATTTGAGACTG	CCCTCCGGCTGCCTGCCTCT	1357
Db	25228	CAGGCCACCTCT	CTGTGCCAAAAGCCCCCTCATTTGAGACTG	CCCTCCGGCTGCCTGCCTCT	25169
QY	1358	GGCACCCACTT	GGCCCTCGGCACAGCCACTCCCTGGAAACATGACT	CTCCGAGCACCCCC	1417
Db	25168	GG-AACCACTT	GGCCCTCGGCACAGCCACTCCCTGGAAACATGACT	CTCCGAGCACCCCC	25110
QY	1418	CGCCCTCTCTCG	AAGGCCCTGCGGTGAGGCACACGCACTGCCTT	CAGCCCCCTCCGGGGGG	1477
Db	25109	CG-CCCTCTCTCG	AAGGCCCTGCGGTGAGGCACACGCACTGCCTT	CAGCCCCCTCC--GGCC	25053
QY	1478	GCCCTTATCAGG	GCATNGGGGCACCTCAGGGCTCCAGAGCAGCTTCA	TCACTACCTGGTGGC	1537
Db	25052	CCCCCTATCAGG	GCATNGGGGCA-CCTCAGGGCTCCAAGCAGCTTCCA	TTCACCTGGTGGC	24994
QY	1538	CACCCAGGCATCT	GCAGCCAGAGAGGCCATCCCCGGACAGCCCTT	GGGGGCAGCCAGCC	1597
Db	24993	CACCCAGGCATCT	GCAGCCAGAGAGGCCATCCCCGGACAGCCCTT	GGGGGAGGCCAGCC	24934
QY	1598	CTTTTCTGCAC	CCCAAGCAGGGTTCTGCCCCCAGAGAGGGTG	CAGCCCCCAACCCAGCA	1657
Db	24933	CTTTTCTGCAC	CCCAAGCAGGGTTCTGCCCCCAGAGAGGGTG	CAGCCCCCAACCCAGCA	24876
QY	1658	GTTCGCCCATG	CCCTCTCTGGCTCTTCCCTCCAGGATCTTCCAAGAG	GGCCCCCTTAGTA	1717
Db	24875	GTTCGCCCATG	CCCTCTCTGGCTCTTCCCTCCAGGATCTTCCAAGAG	GGCCCCCTTAGTA	24816
QY	1718	CCCTAAGCCCTT	TTTGGGACAGCCCCCAGGCA	CCCCCTGCCCTTGC	1777
Db	24815	CCCTAAG-CCCTT	CTTTGGGACAGCCCCCAGGCA	CCCCCTGCCCTTGC	24757
QY	1778	CATTGGACTTA	AGATGGGCGCTGGAGCATCTCTTCTCTGGGAG	GCACAAACCCGGC	1837
Db	24756	CATTGGACTTA	AGATGGGCGCTGGAGCATCTCTTCTCTGGGAG	GCACAAACCCGGC	24697
QY	1838	CCCTGCAGTTT	CCCCAGGGTTCAGCCTCC	CAGCGGAGCTCTTCCCAAGT	1897

Db	24696	CCCTGCAGTCTCCCGAGGGTCAGCCTCCAGCGGAGCTCTTCCCAAGTGAGCTCCCTCAGG	24637
Qy	1998	GTGGGCTCCTCCCAAGTGGGACACAGAGCCTGGCCCCCTCCCTGATGCGGAGGGCTGGACC	1957
Db	24636	GTGGGCTCCTCCAGGTGGGACACAGAGCCTGGCCCCCTCCCTGATGCGGAGGGCTGGACC	24577
Qy	1958	CAGGAGGCTGAGGATCTGTCCGAGCTCCACACCCACCTTGCAGCGGCTTCAGGAAACAGGCG	2017
Db	24576	CAGGAGGCTGAGGATCTGTCCGAGCTCCACACCCACCTTGCAGCGGCTTCAGGAAACAGGTG	24517
Qy	2018	ACCATGCGCAAGTTCTCCCTGGGTGGTGGCGGGGGCTACGCGAGGCGTGGCTGGCTATGGC	2077
Db	24516	ACCATGCGCAAGTTCTCCCTGGGTGGTGGCGGGGGCTACGCGAGGCGTGGCTGGCTATGGC	24457
Qy	2078	ACCTTTGGCTTTGGTGGAGATGCAAGGGGGCATGTGGGGGCAAGGGCCCATGTGGGGCCAGG	2137
Db	24456	ACCTTTGGCTTTGGTGGAGATGCAAGGGGGCATGTGGGGGGCATGTGGGGGGCCCATGTGGGGCCAGG	24397
Qy	2138	ATAGCTGGGGCTGTGTCCAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCC	2197
Db	24396	ATAGCTGGGGCTGTGTCCAGTCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGTCC	24337
Qy	2198	CAGTCGAGGAGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2257
Db	24336	CAGTCGAGGAGCAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	24277
Qy	2258	CCTGTGCTGAGGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2317
Db	24276	CCTGTGCTGAGGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	24217
Qy	2318	GACATCGGGAGGCTCCCTGGTGCAGATCCGGGAGCCTGTGAGGTGATGCGGAGGCGGCC	2377
Db	24216	GACATCGGGAGGCTCCCTGGTGCAGATCCGGGAGCCTGTGAGGTGATGCGGAGGCGGCC	24157
Qy	2378	GACACAAATATCCCTGGACATTTCCGAGGTGGACCCCGCTACCTCAACCTCTCAGACCTG	2437
Db	24156	GACACAAATATCCCTGGACATTTCCGAGGTGGACCCCGCTACCTCAACCTCTCAGACCTG	24097
Qy	2438	TACGATATCAAGTACCTCCCATTCGAGTTTATGATCTTTCAGGAAAGTCCCCAAGTCCGCT	2497
Db	24096	TACGATATCAAGTACCTCCCATTCGAGTTTATGATCTTTCAGGAAAGTCCCCAAGTCCGCT	24037
Qy	2498	CAGCCAGAGCGCCTCCCTCCCATTCGCTGAGGAGGAGCTGGCGGAGTTCCTCCGAGGCCACG	2557
Db	24036	CAGCCAGAGCGCCTCCCTCCCATTCGCTGAGGAGGAGCTGGCGGAGTTCCTCCGAGGCCACG	23977
Qy	2558	TGGCCCTGGCCAGGTGAACCTGGGCCCCCACCAGGAGGAGCTGGAGATCACAAGAGGAGTCAGAG	2617
Db	23976	TGGCCCTGGCCAGGTGAACCTGGGCCCCCACCAGGAGGAGCTGGAGATCACAAGAGGAGTCAGAG	23917
Qy	2618	GATGTGAGACGCGCTGTGGCAGAGGCTGCCCTGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2677
Db	23916	GATGTGAGACGCGCTGTGGCAGAGGCTGCCCTGGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG	23857
Qy	2678	TCAGCGAGCCTTCCCTCCCTGGGAGGAGCCTGGCGCTGGATGAGCCTGCAGAGGCTG	2737
Db	23856	TCAGCGAGCCTTCCCTCCCTGGGAGGAGCCTGGCGCTGGATGAGCCTGCAGAGGCTG	23797
Qy	2738	GGGCTGCGTGAGAGAGTGAAGGCGCTCCGCTGAGGAGCAGATCTCCCGGATCCTGAAAGGGCAGG	2797
Db	23796	GGGCTGCGTGAGAGAGTGAAGGCGCTCCGCTGAGGAGCAGATCTCCCGGATCCTGAAAGGGCAGG	23737
Qy	2798	CCGGAAGGT 2806	
Db	23736	CCGGAAGGT 23728	

Search completed: March 21, 2005, 05:35:32
Job time : 13553.8 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:24:10 ; Search time 4515.62 Seconds
(without alignments)
10741.310 Million cell updates/sec

Title: US-10-077-130-6_COPY_2000_3000

Perfect score: 1001

Sequence: 1 cgcctaccatccggagggtg.....ctgagcgaggggccaatgcc 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_on.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	20435	9 HSAJ2535	AJ002535 Homo sapi
2	830.8	89.0	2768	6 ARS42178	ARS42178 Sequence
3	531.2	53.1	18524	6 CQ730656	CQ730656 Sequence
4	284.8	28.5	24545	9 HSA314898	AJ314898 Homo sapi
5	284.8	28.5	135964	9 AL335393	AL335393 Human DNA
6	229.6	22.9	19475	9 HSA314900	AJ314900 Homo sapi
7	225.4	22.5	260998	2 AC099089	AC099089 Rattus no
8	224	22.4	2488	6 BD160445	BD160445 Primer fo
9	224	22.4	2488	6 AX883644	AX883644 Sequence
10	224	22.4	2488	9 AK024186	AK024186 Homo sapi
11	224	22.4	174612	2 AC023889	AC023889 Homo sapi
12	221.6	22.1	3956	6 C0851220	C0851220 Sequence
13	220.6	22.0	3956	9 AK128447	AK128447 Homo sapi
14	220.2	22.0	244553	2 AC098133	AC098133 Rattus no
15	220	22.0	6400	9 HSA314903	AJ314903 Homo sapi
16	218.8	21.9	211829	10 AL645854	AL645854 Mouse DNA
17	217.6	21.7	10091	9 HSA314901	AJ314901 Homo sapi
18	213.8	21.4	44358	2 AC142478	AC142478 Rattus no
19	212.2	21.2	164766	2 AC026657	AC026657 Homo sapi

C	20	186.6	18.6	250087	2	AC096931	AC096931 Rattus no
C	21	151.8	15.2	526	11	BV077840	BV077840 S208P6660
	22	112.8	11.3	5783	6	CQ715191	CQ715191 Sequence
	23	98.4	9.8	724	6	BD150413	BD150413 Primer fo
	24	98.4	9.8	724	6	AX870351	AX870351 Sequence
	25	98.4	9.8	2184	6	BD160248	BD160248 Primer fo
	26	98.4	9.8	2184	6	AX883309	AX883309 Sequence
	27	98.4	9.8	2184	9	AK023854	AK023854 Homo sapi
	28	93	9.3	5314	10	AK122340	AK122340 Mus muscu
	29	92.4	9.2	5382	9	AB014557	AB014557 Homo sapi
C	30	91.2	9.1	168144	2	CR792456	CR792456 Danio rer
	31	90.8	9.1	5378	6	CQ875354	CQ875354 Sequence
	32	90	9.0	1704	9	AF035292	AF035292 Homo sapi
	33	88.4	8.8	2025	9	AK025946	AK025946 Homo sapi
	34	88.4	8.8	3316	9	BC007201	BC007201 Homo sapi
	35	87.8	8.8	2761	9	BC043916	BC043916 Homo sapi
	36	86.2	8.6	2534	6	AX834822	AX834822 Sequence
	37	86.2	8.6	2534	9	AK097489	AK097489 Homo sapi
	38	84.2	8.4	3015	10	BC056205	BC056205 Mus muscu
	39	81.4	8.1	933	9	BC061909	BC061909 Homo sapi
	40	75	7.5	2417	10	BC051643	BC051643 Mus muscu
C	41	70	7.0	200465	2	AC148786	AC148786 Otolemur
C	42	70	7.0	215101	2	AC148847	AC148847 Otolemur
	43	69.2	6.9	125020	9	AF429315	AF429315 Homo sapi
C	44	66.8	6.7	125020	9	AF429315	AF429315 Homo sapi
C	45	66.6	6.7	193064	2	AC149860	AC149860 Papio anu

ALIGNMENTS

RESULT 1	HSJAJ2535	Homo sapiens mRNA for obscurin (OBSCN gene).	20435 bp	mRNA	linear	PRI 14-SEP-2001
LOCUS	HSJAJ2535	Homo sapiens mRNA for obscurin (OBSCN gene).				
DEFINITION	AJ002535					
ACCESSION	AJ002535.1	GI:15026973				
VERSION	OBSCN gene; obscurin.					
KEYWORDS	OBSCN gene; obscurin.					
SOURCE	Homo sapiens					
ORGANISM	Homo sapiens					
REFERENCE	1	Young, P., Ehler, E. and Gautel, M.				
AUTHORS	1	Young, P., Ehler, E. and Gautel, M.				
TITLE	1	Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor protein involved in sarcomere assembly				
JOURNAL	1	J. Cell Biol. 154 (1), 123-136 (2001)				
MEDLINE	21342081					
PMID	11448995					
REFERENCE	2	(bases 1 to 20435)				
AUTHORS	Gautel, M.S.					
TITLE	Direct Submission					
JOURNAL	Submitted (31-OCT-1997)	Gautel M.S., Structural Biology Division, European Molecular Biology Laboratory, Meierhofstr. 1, Heidelberg, 69117, GERMANY				
REMARK	Revised by author 20-JUL-2001					
FEATURES	Location/Qualifiers					
source	1..20435					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/chromosome="1"					
	/map="1q42.1"					
	/cell_type="cardiomyocyte"					
	/tissue_type="cardiac"					
	/dev_stage="adult"					
gene	1..20435					
	/gene="OBSCN"					
CDS	72..19934					
	/gene="OBSCN"					
	/function="titin binding"					
	/codon_start=1					
	/evidence=experimental					


```

/product="obscurin"
/protein_id="CAC4768.1"
/db_xref="GI:15026974"
/db_xref="GOA:O96AA2"
/db_xref="UniProt/TREMBL:O96AA2"
/translation="MDQPFSCAPFLRPKAFVVSUGKDATLSQIVGNTPPQVSWE
KQQPVTAGPRLAQDGLYKLTLLDALGDSGGYVCRARNAIGEAFAVGLQVDAB
AACQAPAFHLPRLTISIRREGSEATFRGQSGPRPAWSKODRGLBGPDPGRVIR
BELGASALIRIARAPRODGTVEVRAENPLGAASAAALVVDSDAAPTSPGTAA
LHAIQRRREARAGAPGASPSTRTCTTEGKHARLSCVVTGEPKPTVMKKDGO
LVTEGRHVVEKESATLFCQSDRGLYCTASNLVGQYSSVVLVVRPAVP
PKRLQDLVEKESATFLCEVPQPSSTEAAMKFEKTRLWASAKYGBIEGKERLTIVR
NYSADDDAYICETPEGSKTVAELVAGNLLKRLPKRTAVAGTMAPFCVBLAVPVG
VHMLRQEVVAGRVASIASGCTRHTLTISQCLEDDQVAFMAGDCQTSFRFCVSAP
RKPLQPPVPVKARMESSVILSWPPPHGPERPVTIDGYLVEKKKLGTWYRICHFA
EWATPELTVDVAEGNFOFRVSAINSGFQSPYLPBPGTTHLAPKLAVRPLKAVOA
VEGGRVTSVDLTVASACEMFLDGOALKASSVYRTHCDRTBHTLTIREVPSLHGAQL
KFAVGISSIRMEVRAAPGLTANKPAPAAAAREVLARLHESAQLLAELSQAQAAWTML
KDRTLSCPDTEVQASAGRVLLRDVARDADAGLYECVSRGRIAYOLSVQGLARFL
KDMWAGSCVAVAGPAQFECETSEAHVHVHMDGMBELGHSGERFLQEDVGTGRLHV
AATVTRQDEGYSICRVGSDSDVFLRVSEPKVPFAKEQLARRKLQAAGASATLSCEV
AQAOETVWYKDGKLSKSSKVMCEATGCTRLRVVQOAGQADAGEYSCAAGORLSFH
LDVKEPKVFAKDVVAHSEVOAEAGANATLSCEVAQAQAEVWYKDGKLSKSSLKVHV
EAKGCRRLVQQAGKTADGYSCEARGQSVFRLHYTEPKMFAKQSVNHEVQABA
CASAMLSCEVAQOETVWYKDGKLSKSSKVMGKVGCTRLRVLPQKADAGEYSC
BAGGQSVFHLHIIEPKGVFAKEQSVNHEVQAEAGTAMLSCEVAQOETVWYKDGK
KLSSSKVMGKVGCTRLRVVQOQKADAGEYSCAAGQSVFQHLHIEPKAVFAKEQ
LVHNEVTEAGASATLSCEVAQOETVWYKDGKLSKSSKVRTEAAGCMROLVVQQA
QOADAGEYTCAGGQRLSFHLDVSEPKAVFAKEOLAHKVKQAEAGATLSCEVAQA
TEVWYKDGKLSKSSKVMGAGCTRLRVVQOQADTGEYSCAAGQRLSFSLDVA
EPKVVFAKEQPVHREVAQASATLSCEVAQOETVWYKDGKLSKSSKVMGAEAGT
CTRLRVVQOQADAGEYSCAAGQRLSFHLDVSEPKAVFAKEQAEVWYKDGKLSKSSK
TLSCEVAQOETVWYKDGKLSKSSKVMGAGCTRLRVVQOQADAGEYSCAAGQ
QRLSFHLDVSEPKAVFAKEQPAHREVQAEAGASATLSCEVAQAQOETVWYKDGKLS
SKVRVEAGCTGCTRLRVVQOQADAGEYSCAAGQRLSFRLHVAELBQIIEPKVCE
PLVKEHEDIILTATLTATSAATVWLDKGVEIRSRKHETASQGDTHLTIVHGAQVL
DSALYSCRVAEGQDFPVQVEEAAKFCRLLEPVGSELGGTATLACLSFACAEVWVR
CGNTQPRVGRFQWGAEPVRLVLGLRAEDAGEYCESDRHTLSAQLTVSPRVYVK
FMGSLGTVVAEGEAGTFCVSPSDVAVMFDRGALLQSPSEKFAISQSASHSLSITS
DLWRNATQALTVRAEGASSAAELVREAPLVFKKLEPQTVBERSSVDTLVELTRPWP
ELRTRNATQALTVRAEGASSAAELVREAPLVFKKLEPQTVBERSSVDTLVELTRPWP
RQVLVRGLQAEAREOGATVMEVQLSHADVDGSWTRDGLRFQCGPTCHLAVRGMHT
LTLGSLRPDSGLMFKABGQTSARLVTVLSEFYSRQDQVVTTEKEKLTCELS
RPNVDVRLKDGVELRGAQMAIAAGACHSLTIVYCEFAFDQVQADHAQSSASV
KQVQRTIYLRVLAEDAGEIQFVAENAESRAQLVKELPVLTVLRPLRDKIAEKHR
GVLCEQVSRASQVWRFKGSQELQPGPKYELVDGLYKRLIISDVHAEDBDTVDQAG
DVKTSADQFVBEQSIITTVRGLODVTVMEPAPWPECETSIPSVRPPKWLKGTVLQAG
GNLGLQEGTVHRLMARTCTMTGTVHFTVGSKRSARLVVSDIPVWLRPLPXTG
RELOSVLSCDFRPAKAVQYKDDTFLPSFKFMSLQGMALRLILRLMPADAGVY
RCQGAHSSTEVTVAREVTVTGLQDAEATEEGWASFCELSHEDVEEWSLNGMP
LYNDSFHEI SHKRRHTLVLSIQORADAGIVRASSLKVSTSVARLEVRKVPVHLKALD
DLSABERGTLALQCEVSDPEAHVVRKDGVLQSPDKYDFLHTAGTGLVHVDVSPED
AGLTVCHVGEETRARVRVHDLHVGTIKRLKTMVLRGESSPECVLSHESADPAMW
TVGGTKVSSSRFOATQGRKYILAVREAPSDAGEVVFVSRGLTSKASLIVRPA
IIFLEPDQWVAPGEDVLRCELBRAGTPVHMLKDKRAIKRSQKTDVVCBGTMAWLVIR
GASLKAGEYTCVEASKTSASLHVEEKANCFTEELNLQVEEKGTAVFTCKTEHPAA
TVWTKGLLELRAGKQSQBGLTLTLTISALEKADSDYTTCDIGQASRAQLLVQV
RRVHIIDBLEDDVDQEGSSATFRCRIISPANVEPWHPLDTPHLPLHAINELIDAQPGY
HVLTLROLALKDGTIYFEAGDORASALRVTEKPSVFSRELTDATITEGDI TLVCE
TSTCDIPMCWTQKGLRGSARCQSHGHEHQALLITGATLQDSGRYKCBAGACASS
IVRVHARPVRQBAKLDLEVEGANTLRCVLSVAAPKWCYNGYVNPFGDKYLSURQ
EGAMLEVLNRLRPQDSGRYSFGDQDTTATLTVTALPAQFI GKLRNKEATEGATAT
LRCELSKTAPEVRKGETLDRDGYCLRQDGAMCELOIRGLAMVAAEYSCVGBER
TSASLTIDPMPAHP IGRHREQSEGTATLRCELSKAAPEVRKGRSRLRDRGSHSL
RODQACELOICGLAVADAGEYSCVCEERTSATLTVKLPKAFTEGLRNEBVEGAT
AMLMWCELSKVAPEVRKGPENLRDGRVILROBGTCELOICGLMADAGEYLCVCGQ
ERTSATLITRALPARFIVLVDNQAREGATVLCHELNSAPVWFKRNGSETLRDGDRY
SLRQDGTKELOIRGLANADTGEYSCVCGQERTSAMLTVRALPIKTETGLRNEBTEG
ATAVLRCELSKMAPEVRKGETLDRDGRSLRQDGARCELOIRGLVAEDAGEYLCMC
GKERTSAMLTVRAMPKFI EGRNNEATEGDTATWCELSKAAPEVRKGRHETLRDGD
RHSRLRQDSRCELOIRGLAVVDAGEYSCVCGQERTSATLTVRALPARFIVDKNQBAR
EGATAVLQCELSKAAPEVRKGETLDRDGRYSLRQDGTKELOIRGLSVADTGEYSC
VCQQERTSATLTVRAQPVPRFPLQSLQAEEGSATLQCELSSEPTATVWMSKGLQQLQ

```

ORIGIN

Query Match	100.0%	Score 1001;	DB 9;	Length 20435;
Best Local Similarity	100.0%;	Pred. No. 1.1e-158;		
Matches 1001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	CGCTACCATCCGGGAGGTGCCCGCAGCCTGACCGGGGCGCAGCTGAAAGTTCGTGCCCA	60	
Db	2071	CGCTACCATCCGGGAGGTGCCCGCAGCCTGACCGGGGCGCAGCTGAAAGTTCGTGCCCA	2130	
Qy	61	ACGGCATTGAGACACATCCGATCGAGTCCGGCGGCGCCCGAGGCTGACTGCCAACA	120	
Db	2131	ACGGCATTGAGACACATCCGATCGAGTCCGGCGGCGCCCGAGGCTGACTGCCAACA	2190	
Qy	121	AGCCGCGCAGCGCAGCTGCCCGGAGAGTGTGGCTGCGCTGCAACGAGGAGCGCAGCTGC	180	
Db	2191	AGCCGCGCAGCGCAGCTGCCCGGAGAGTGTGGCTGCGCTGCAACGAGGAGCGCAGCTGC	2250	
Qy	181	TGGCTGAGCTGTGATCAGGCTGCGGCTGTGAGCTGGCTGAAGGATGGTCGACACTGT	240	
Db	2251	TGGCTGAGCTGTGATCAGGCTGCGGCTGTGAGCTGGCTGAAGGATGGTCGACACTGT	2310	
Qy	241	CCCCAGGCCCCCAAGTATAGGTGACGGCATCGCGGGGCGGGGTGCTCTTGTGCGAG	300	
Db	2311	CCCCAGGCCCCCAAGTATAGGTGACGGCATCGCGGGGCGGGGTGCTCTTGTGCGAG	2370	

```

ANGREPRLOQCTAELVLQDLQREDTGEYTCCTCGSQATSATLTVTAAFPVRELQHQ
EVDGEGTHLCCSLRASAGVSKSGSLQLPPCAKYQMVODGAAAYELVRCVBOEDAG
DCTDCTGHTOSMASLSVVRPBPKTKRLOSLOEQTGDIARLCCOLSDAESAVQWLK
EGVELHAGPKYKRSQCATRELLIHLEAKDTEGYACVTGGQKTAASLRVTEPEVTIV
RGLVDAETADEDEVFSCVSRAGATGVQMCLOPLQPSNEVEVAVRDRGRIHRLRLK
FTAPEDAGTVYFHLGNHASSAQLTVRAPEVTILEPLQVLSQEGDASFCQRLSRASG
QEARWALGVPLQANEMNDITVEQGTIHLHTLKHVTLEDAGTVSHFVGTCSSEALQKV
TAKNTVVRGLNVEALRGBALFCOLSQPEVAHATWLLDDDEPVRTSENARVFPENG
LRLHLLKLNLRPODSRCRVTLAGDMVTSLELVRGWLELLEPLKNAARVAGAARPT
CTLSEAVPVGEASWYINGAAVQPDSDWTVTGDSHQALLRISQPHHAGEVTFACRD
AVASARUTVLGLPDPPEDAEVAHSSHTVTLSWAAPMSDGGGLCGVRCVKEGATGQ
WRILCHELVPFECVVDGLAPGETVFRVAAPVPGAGEPVHLPTQVRLAEPFKPVPVPG
PASPESQVAAAGEDVLSLEVAEAGEVIWHKGMERIQPGFERFEVVGQGMVLVIG
FTADQEGYHCLGAQSGICPAAATFOVALPSASVDEAPQSPLEPAEAQEGDHLHLEWA
LAKRMRSPETPLDTSISELPEEDGRSQRLOPQEAEEVAPDLSEGYSTADELARTGDADL
SHTSDDSEAGTPESLVYLKAGRPCTSLASKVGAAPAASVKPQOQQPFLAARPP
LGLDLSKDLGDPDNDKAAVKIQAFKGKVKRKKMQQEGPMFSHTPQTEAGVQDALAR
LECVAASDARWRLKDGVELTDRHHHLDQLGDGTCSLIAGLDRADAGCYTCQVS
NKFGQVTHSACVVVSGSEASESSSGEGLDDAFRRARRLRLHRLFKSPAIEVDEELF
LSADEGPAEPPEPADWOTYREDSHFCIRFEALTTEARQAVTRFOEMFALIGIVEIKL
VEQPRRVEVICIKETPAPVVPPEPLPSLLTSDAAPVLTQLQNEQVQDGYPVSVFDCV
VTGQPMPSVWFMDKGLLEDDHVMNEDQGGHQLIITAVPDMGTVYRCLANSWG
VSTKAELRVDLTSTDYTDADATESSYFSAQYLSRQEGTESTIDSGQLPVVE
ELRDLQVAPQTRLAKFLQKVGYPAPRLVWFKDQGPLTASAHHEQLVPPRMLRLEIFK
VTRDSSGQAAAYISNAGAAYSARLLVRGPDEPEKPSADSVHHEQLVPPRMLRLEIFK
VGRHOCTYTICIASNAAGQALCSALHVSGLPKVBEQKVKEALISTFLQOQTQALISA
QGPOTASFADLQGRKEEPLAAKEALGHLISAEVGTTEEFKLKLTQITENWSAKITQA
KLQVPGDSDLEDKSTPSAPRHSRPPSSIQBSSSESDGDARGEIIFIIVVTADLV
PLGQAQDAILTREGQVVLDDAAHPLVRLTKSSPSRSGQWGPAYLDRKLKLP
EWGAAPPEPFGAEVSEDEYKARLSVIOQLLSEQAFBELOFLASHHLQELRCPH
VPTAVAGQKAVIPRNVRDIGHRSHFSFLQELQCDTDDVMACFIKQAAEQYLEFV
ELIRNARVQNCALLQEAQVAVSALPQRAENKLVSLMENVPTQLQALGEPVIRQGHF
IWEAGAPKMPKGNHNVLPRLNHLVCTPRDRSDTDTVSIVFRNMKLSLIDLND
QVEDDRAEFVEQERDSVRKYLQARTAIKSSWYKEICGIQORLALFVRPPDFEE
ELQADAEFVETKLCRCVTGPKPVI SWYKDGKAVQDPPHILIIDPDGSCALILDS
LTCVQSGCYMFAASANGCSTLKVLPVPPVFNKVLASPVBEGBDAQCTIEGA
PYQIWRYKDGALLTNGKFTQSLSEPSGLLVIVIRAAKEDIGLVECELVNRLGSAR
ASALRIQSPMLQAOQCHREOLVAEVTEOETKVPKKTVIIIEETITTVVKSQRGR
RSPSKSPSPSPSCSAPLRIGLLAILVLGAGQPPRPAEPFGKVPVPTLYGTATP
EAHSPALPGLSGGQPKWVEVEETIEVRKMGQGVSPTEVESRSGHLLFTLPATP
GDPNSENNNKLLQAQWAOQTAMVVRVREPLFVDRGSDVMDAASGMGSLEBETHC
EAGBEEGDGDAFTVEESQDTHSLGDRDPKILTHNGRMLTLADLEDYVVEGEETHC
GGPGPADPPCEVSVIOREIGEPTVGOPVLLSVGHALCPRGPLGFRPEPGASPP
GPVRSLEGTSPFLREAPRVGSAPWTSQFCTRIRRSADSGSSFTTELSTQTQVNF
TVGETVLIHCPRDDDEAAQp"

```


Qy	301	ATGTGGCCCGGGACGATGCAAGGCCTCTTACGAGTGCCTACGCGCGGGGGCCGCATCGCCT	360
Db	2371	ATGTGGCCCGGGACGATGCAAGGCCTCTTACGAGTGCCTACGCGCGGGGGCCGCATCGCCT	2430
Qy	361	ACCAGCTCTCGGTGCARAAGGCTTCGCGCGCTTTCTGCACAAGGACATGGCGGGCAGCTGTG	420
Db	2431	ACCAGCTCTCCGTGCARAAGGCTTCGCGCGCTTTCTGCACAAGGACATGGCGGGCAGCTGTG	2490
Qy	421	TGGATCGCTGGCTTGGGGGGCCCGCGCGAGTTTGAGTGTGAGACTCTCCGAAGCCACAGTCC	480
Db	2491	TGGATCGCTGGCTTGGGGGGCCCGCGCGAGTTTGAGTGTGAGACTCTCCGAAGCCACAGTCC	2550
Qy	481	ACGTGCACCTGGTACAAGGATGGCAGTGGGCCACTCCGGTGAGCGCTTCTTGACAGG	540
Db	2551	ACGTGCACCTGGTACAAGGATGGCAGTGGGCCACTCCGGTGAGCGCTTCTTGACAGG	2610
Qy	541	AGGATGTGGGGACCGCGGCAACCGGCTGCTGGCAGGCCACAGTCCACGAGCAGGATGAAGGCA	600
Db	2611	AGGATGTGGGGACCGCGGCAACCGGCTGCTGGCAGGCCACAGTCCACGAGCAGGATGAAGGCA	2670
Qy	601	CCTACTCCTCGCGGTGGGCGAGGACTCTGTGAGACTTCCGGCTCCCGGCTCTCTGAGGCCA	660
Db	2671	CCTACTCCTCGCGGTGGGCGAGGACTCTGTGAGACTTCCGGCTCCCGGCTCTCTGAGGCCA	2730
Qy	661	AGGTGGTGTTTGCTTAAGGAGCAGCTGGCAGCGCAGGAAAGCTGCAGGCAGAGCGCAGGAGCCA	720
Db	2731	AGGTGGTGTTTGCTTAAGGAGCAGCTGGCAGCGCAGGAAAGCTGCAGGCAGAGCGCAGGAGCCA	2790
Qy	721	GTGCCACACTGAGCTGCGAGGTGGCCACAGGCCACAGCGAGGTGACGTGTGTACAAGGATG	780
Db	2791	GTGCCACACTGAGCTGCGAGGTGGCCACAGGCCACAGCGAGGTGACGTGTGTACAAGGATG	2850
Qy	781	GGAGAAGCTGAGCTCCAGCTCGAAGTGTGTCATGGAGGCCACAGGCTGCACGGGCAGGC	840
Db	2851	GGAGAAGCTGAGCTCCAGCTCGAAGTGTGTCATGGAGGCCACAGGCTGCACGGGCAGGC	2910
Qy	841	TGGTTGTGCAGCAGGCGCAGCGGATGTCGGGGGAGTATAGCTGCGAGGCTTGGGGGCC	900
Db	2911	TGGTTGTGCAGCAGGCGCAGCGGATGTCGGGGGAGTATAGCTGCGAGGCTTGGGGGCC	2970
Qy	901	AGCGGCTCTCCTTCCATCTGGATGTCAAAAGAGCCCAAGGTGGTGTTTGCCAAGGACGAGG	960
Db	2971	AGCGGCTCTCCTTCCATCTGGATGTCAAAAGAGCCCAAGGTGGTGTTTGCCAAGGACGAGG	3030
Qy	961	TGGCACACAGTGAGGTGCGGCTGAGCGAGGGGCCAATGCC	1001
Db	3031	TGGCACACAGTGAGGTGCGGCTGAGCGAGGGGCCAATGCC	3071

RESULT 2			
ARS542178			
LOCUS	ARS542178	2768 bp	DNA
DEFINITION	Sequence 426 from patent US 6743619.	linear	PAT 08-OCT-2004

ORGANISM: Unknown;
Unclassified.
REFERENCE 1 (bases 1 to 2768)
AUTHORS Tang,Y.T., Zhou,P., Goodrich,R., Liu,C., Asundi,V., Ren,F., Zhang,J., Zhao,Q.A., Yang,Y., Xue,A.J., Wehrman,T., Wang,J.-R., Wang,D. and Dmanac,R.T.
TITLE Nucleic acids and polypeptides
JOURNAL Patent: US 6743619-A 426 01-JUN-2004;

```

ORIGIN
Query Match      83.0%; Score 830.8; DB 6; Length 2768;
Best Local Similarity 95.3%; Pred. No. 5.6e-130;
/organism="unknown"
/mol_type="genomic DNA"

```

Matches	856;	Conservative	0;	Mismatches	42;	Indels	0;	Gaps	0;
QY	1	CGCTCACCATCCGGAGAGTGTCCCGCCAGCTCTGCACGGGGCGCAGCTGAAGTTCGTGGCCA	60						
DB	1861	CGCTCA CCAATCCGGAGAGTGTCCCGCCAGCTCTGCACGGGGCGCAGCTGAAGTTCGTGGCCA	1920						
QY	61	ACGGCAATTGAGAGCAGCATCCGATGAGGTCTGGGGCGGCCCAAGGCTGACTGCGCCAACA	120						
DB	1921	ACGGCAATTGAGAGCAGCATCCGATGAGGTCTGGGGCGGCCCAAGGCTGACTGCGCCAACA	1980						
QY	121	AGCGCCAGCCGAGCTCCCGGAGAGTCTGGCTGGCTGGCGTGCACGAGGAGGCGCAGCTGC	180						
DB	1981	AGCGCCAGCCGAGCTCCCGGAGAGTCTGGCTGGCTGGCGTGCACGAGGAGGCGCAGCTGC	2040						
QY	181	TGGCTGAGCTGTTCAGATCAGGCTCGCGCTGTGACGTGGCTGAAAGGATGSGTCGCACACTGT	240						
DB	2041	TGGCTGAGCTGTTCAGATCAGGCTCGCGCTGTGACGTGGCTGAAAGGATGSGTCGCACACTGT	2100						
QY	241	CCCAGAGCCCAAGTATGAGGTGAGGCATCGCGCGGCGCGGGTGTCTCTTTGTGGCAG	300						
DB	2101	CCCAGAGCCCAAGTATGAGGTGAGGCATCGCGCGGCGCGGGTGTCTCTTTGTGGCAG	2160						
QY	301	ATGTGGCCCGGAGCATGCAAGGCTCTACAGATGCGTCACGCGGGGGCGCGATCGCCT	360						
DB	2161	ATGTGGCCCGGAGCATGCAAGGCTCTACAGATGCGTCACGCGGGGGCGCGATCGCCT	2220						
QY	361	ACCAAGCTCTCGTCAAGAGGCTCCGCGCTTTCTGCACAAGAGCATCGCGGGCAGCTGTG	420						
DB	2221	ACCAAGCTCTCGTCAAGAGGCTCCGCGCTTTCTGCACAAGAGCATCGCGGGCAGCTGTG	2280						
QY	421	TGGATGCCGTGGTCTGGGGGCCGCGCAGTTTGAAGTGTGAGCTCTCGAAGCCACGCTCC	480						
DB	2281	TGGATGCCGTGGTCTGGGGGCCGCGCAGTTTGAAGTGTGAGCTCTCGAAGCCACGCTCC	2340						
QY	481	ACGTGCACGTGTACAAGATGGCATGAGCTGGGCCACTCGGTGAGCGCTTCTTGACGG	540						
DB	2341	ACGTGCACGTGTACAAGATGGCATGAGCTGGGCCACTCGGTGAGCGCTTCTTGACGG	2400						
QY	541	AGGATGTGGGACCGCGCACA CCGGCTGTGGCAGCCACAGTCA CAGAGCAGGATGAAGCCA	600						
DB	2401	AGGATGTGGGACCGCGCACA CCGGCTGTGGCAGCCACAGTCA CCAAGCAGGATGAAGCCA	2460						
QY	601	CCTACTCTCTGCCGCTGGCGAGCACTCTGTGCACTTTCGGCTCCGGCTCTCTGAGGCCA	660						
DB	2461	CCTACTCTCTGCCGCTGGCGAGCACTCTGTGCACTTTCGGCTCCGGCTCTCTGAGGCCA	2520						
QY	661	AGGTGGTGTTCGTTAAAGACAGCTGCACA CCGCAGGAGCTGCAGGCCAGAGCGCAGGCCA	720						
DB	2521	AGGTGGTGTTCGTTAAAGACAGCTGCACA CCGCAGGAGCTGCAGGCCAGAGCGCAGGCCA	2580						
QY	721	GTGCCACACTGAGCTGCGAGGTGGCCCGCAGGCCAGACGAGTGCACCTGGTACAAAGGATG	780						
DB	2581	GTGCCACACTGAGCTGCGAGGTGGCCCGCAGGCCAGACGAGTGCACCTGGTACAAAGGATG	2640						
QY	781	GGAAAGAACTGAGTCTCAGCTCGAAAGTGTGCATATGGAGGCCACAGGCTGCACGCGCAGGC	840						
DB	2641	GGAAAGAACTGAGTCTCAGCTCGAAAGTGCATATGGAGGCCACAGGCTGCACGCGCAGGC	2700						
QY	841	TGGTTGTGCACAGGCAGCCAGCGCATGCCCGGGAGTATACCTGCGAGGCTGGGGG	898						
DB	2701	TGGTTGTGCACAGGCAGCCAGCGCATGCCCGGGAGTATACCTGCGAGGCTGGGGG	2758						

[illegible]

10

REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof
JOURNAL Patent: WO 02068579-A 16590 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1. 18524
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 53.1%; Score 531.2; DB 6; Length 18524;
Best Local Similarity 90.7%; Pred. No. 9.1e-80;
Matches 566; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
QY 378 GGCCTCGCGCGCTTTCTGCACAAAGACATGCGCGCAGCTGTGTGATGCCGTGGCTGGG 437
Db |||||
309 GGCCTCGCGCGCTTTCTGCACAAAGACATGCGCGCAGCTGTGTGATGCCGTGGCTGGG 368
QY 438 GCGCCGCGCGCAGTTTGTAGTGTGAGACCTCCGAAAGCCCAAGCTCCAGCTGCTGTGTACAA 497
Db |||||
369 GCGCCGCGCGCAGTTTGTAGTGTGAGACCTCCGAAAGCCCAAGCTCCAGCTGCTGTGTACAA 428
QY 498 GATGCATGTGAGCTGGCGCACTCCCGTGTGAGCGCTTCTTTCAGAGGAGTGTGGGACCGCG 557
Db |||||
429 GATGCATGTGAGCTGGCGCACTCCCGTGTGAGCGCTTCTTTCAGAGGAGTGTGGGACCGCG 488
QY 558 CACCGCTGTGTGGCGCCACAGTCACAGGAGGATGAAGGACACTCTCCTGCGCGGTG 617
Db |||||
489 CACCGCTGTGTGGCGCCACAGTCACAGGAGGATGAAGGACACTCTCCTGCGCGGTG 548
QY 618 GCGCAGGACTCTGTGGACTTCCGCTCCGCTCTCTGAGCCCAAGGTTGTTTGTCTAAG 677
Db |||||
549 GCGCAGGACTCTGTGGACTTCCGCTCCGCTCTCTGAGCCCAAGGTTGTTTGTCTAAG 608
QY 678 GAGCAGCTGGCAGCAGGAAGCTGCAGGACAGGACGAGGACGCTGCCACACTGAGCTGC 737
Db |||||
609 GAGCAGCAGGCTGCAGGAGGAGTGCAGCTGAGTGGGGCCAGCGCACCTGAGCTGT 668
QY 738 GAGTGGCCAGGCCCAGACGAGGAGTGCAGCTGTGTGTACAGGATGGAAAGAGCTGAGCTCC 797
Db |||||
669 GAGTGGCCAGGCCCAGACGATGAGGTGACGCTGTGTGTACAGGACGGAAGATTGAGCTCC 728
QY 798 AGCTCGAAAGTGTGCATGCGAGCCACAGGCTGCAGCGCAGGCTGTGTGTGCACAGGCA 857
Db |||||
729 AGCTCGAAAGTGTGCATGCGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 788
QY 858 GCGCAGGCGGATGCGCGGAGTATAGCTGCGAGGCTGGGGCCAGCGGCTCTCTCTTCCAT 917
Db |||||
789 GCGCAGGCGAGCTCGGAGAGTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 848
QY 918 CTGATGTCAAAGAGCCCAAGGTTGTTTGTCCAAAGACAGGTTGGCACAGTGAAGTG 977
Db |||||
849 CTGACGTTGTGAGAGCCCAAGGTTGTTTGTCTAAGGAGAGTGTGTGTGTGTGTGTGTGTGTGT 908
QY 978 CAGGCTGAGGCGAGGCGCAATGCC 1001
Db |||||
909 CAGGCGAGGCGAGGACAGTGC 932

RESULT 4
HSA314898
LOCUS HSA314898 24545 bp DNA linear PRI 21-MAY-2002
DEFINITION Homo sapiens partial OBSCN gene for obscurin, exons 5-11.
ACCESSION AJ314898
VERSION AJ314898.1 GI:21104327
KEYWORDS OBSCN gene; obscurin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1
Young, P., Ehler, E. and Gautel, M.
Obscurin, a giant sarcomeric Rho-GEF protein involved in sarcomere assembly
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 24545)
TITLE Gautel, M.S.
JOURNAL Direct Submission
TITLE Submitted (22-MAY-2001) Gautel M.S., Physikalische Biochemie,
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse
11, Dortmund, 44227, GERMANY
FEATURES
source
1. 24545
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
332..10891
/gene="OBSCN"
join(332..592,948..1256,1720..1986,2578..2859,3170..3445,
5481..5750,10616..10891)
/gene="OBSCN"
join(332..592,948..1256,1720..1986,2578..2859,3170..3445,
5481..5750,10616..10891)
/gene="OBSCN"
/codon_start=3
/product="Obscurin"
/protein_id="CAC85746.1"
/db_xref="GI:21104328"
/db_xref="UniProt/TreMBL:OBHN8"
/translation="NLLRLPKTAVRVDGTAMFCVLEAVPYGPVHNLNQBEEVAVGG
RYAISAEGRTHLITISQCCLEVDGVQAFMAGDCQSTQFCVSAPRKPPIQPPVDPVVK
ARMSSVILSWSPFPPGPERPVTIDGLVEKKLGLTYWIRCHEAEVATPELTADV
EENFQFRVSALNSFGQSPYLEFPFPGTVHLAPKLAVRTPLKVAQVAGGSEVTFSDLT
ASAGEVFLDGLQALKASSVYEIHCDRTHLTITREVPASLHGALKEFVANGISIRME
VRAAGPLGTANKPPAAAREVLARLHEEAQLAELSDQAAATWIKDRTLSGPQKVEV
QASAGRVLLVDVARDADGLVECVSRGRIAVQLSVQGLARFLHKDMAGSCVDAVAG
GPAQFECSTSEAHVHVHVYKDGMEIGHGSGERFLQEDVGTGTHRLVAATVTRQEGTYS
RVGESVDRLRVSEPKAVPAKQPAQCREVQAEVGSATLSCEVADQDMEVETWYKDGK
KLSSSKVHEAVGCMRRLLVVOOQADSGEYSCSARGQVSRFLDDEPKVFAKBLA
RRKLQAEAGASATLSCEVAQAQTEVTWYKDGKLLSSSKVCMETGCTCTRLRVVQQAQ
ADAGEYSCSAGGQRLSFHLDFK"
332..592
/gene="OBSCN"
/number=5
593..947
/gene="OBSCN"
/number=5
948..1256
/gene="OBSCN"
/number=6
1257..1719
/gene="OBSCN"
/number=6
1720..1986
/gene="OBSCN"
/number=7
1987..2577
/gene="OBSCN"
/number=7
2578..2859
/gene="OBSCN"
/number=8
2860..3169
/gene="OBSCN"
/number=8
3170..3445
/gene="OBSCN"
/number=9
3446..5480
/gene="OBSCN"
/number=9


```

exon      5481.. 5750
           /gene="OBSCN"
           /number=10
intron    5751.. 10615
           /gene="OBSCN"
           /number=10
exon      10616.. 10891
           /gene="OBSCN"
           /number=11

ORIGIN
Query Match      28.5%; Score 284.8; DB 9; Length 24545;
Best Local Similarity 99.3%; Pred. No. 2.1e-38;
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 92 CCGGGCGGCCCGGCTGACTGCGCAACAAAGCCGCGCAGCTGCCCGGAGGTGCT 151
Db 2573 CCCAGCGGCCCGGCTGACTGCGCAACAAAGCCGCGCAGCTGCCCGGAGGTGCT 2632

Qy 152 GGCTCGGCTGCACGAGGAGGCGCAGCTGCTGGCTGAGCTGTTCAGATCAGGCTCGGCTGT 211
Db 2633 GGCTCGGCTGCACGAGGAGGCGCAGCTGCTGGCTGAGCTGTTCAGATCAGGCTCGGCTGT 2692

Qy 212 GACGTGGCTGAAGATGCTGCACACTGTCCCGAGGCCCAAGTATGAGTGCAGGCATC 271
Db 2693 GACGTGGCTGAAGATGCTGCACACTGTCCCGAGGCCCAAGTATGAGTGCAGGCATC 2752

Qy 272 GCGCGGGCGGGGTGCTCTTGTGCGAGATGTGGCCGCGGACGATGCAGGCCTCTACGA 331
Db 2753 GGCCGGGCGGGGTGCTCTTGTGCGAGATGTGGCCGCGGACGATGCAGGCCTCTACGA 2812

Qy 332 GTGGCTACGCGCGGGGCGCGCTACGCTCTCCGTGCAGG 379
Db 2813 GTGGCTACGCGCGGGGCGCGCTACGCTCTCCGTGCAGG 2860

RESULT 5
AL353593
LOCUS      AL353593      135964 bp      DNA      linear      PRI 13-FEB-2002
DEFINITION Human DNA sequence from clone RP5-1139B12 on chromosome 1q42.1-43,
complete sequence.
ACCESSION AL353593
VERSION   AL353593.33 GI:18673899
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 135964)
AUTHORS   Almeida,J.
TITLE     Direct Submission
JOURNAL   Submitted (13-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
COMMENT   On Feb 14, 2002 this sequence version replaced gi:17977879.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WormPep; Information on the WormPep
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP5-1139B12 is from the library RPCI-5 constructed by the group of
Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP5-1139B12.
Location/Qualifiers
1..135964
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="q42.1-43"
/clone="RP5-1139B12"
/clone_lib="RPCI-5"
25928..26054
/misc_feature
/notes="Sequence from overlapping clone RP11-520H14
(ALJ59510). Assembly confirmed by restriction digest."

ORIGIN
Query Match      28.5%; Score 284.8; DB 9; Length 135964;
Best Local Similarity 99.3%; Pred. No. 1.7e-38;
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 92 CCGGGCGGCCCGGCTGACTGCGCAACAAAGCCGCGCAGCTGCCCGGAGGTGCT 151
Db 3743 CCCAGCGGCCCGGCTGACTGCGCAACAAAGCCGCGCAGCTGCCCGGAGGTGCT 3802

Qy 152 GGCTCGGCTGCACGAGGAGGCGCAGCTGCTGGCTGAGCTGTTCAGATCAGGCTCGGCTGT 211
Db 3803 GGCTCGGCTGCACGAGGAGGCGCAGCTGCTGGCTGAGCTGTTCAGATCAGGCTCGGCTGT 3862

Qy 212 GACGTGGCTGAAGATGCTGCACACTGTCTCCCGAGGCCCAAGTATGAGTGCAGGCATC 271
Db 3863 GACGTGGCTGAAGATGCTGCACACTGTCTCCCGAGGCCCAAGTATGAGTGCAGGCATC 3922

Qy 272 GCGCGGGCGGGGTGCTCTTGTGCGAGATGTGGCCGCGGACGATGCAGGCCTCTACGA 331
Db 3923 GGCCGGGCGGGGTGCTCTTGTGCGAGATGTGGCCGCGGACGATGCAGGCCTCTACGA 3982

Qy 332 GTGGCTACGCGCGGGGCGCGCTACGCTCTCCGTGCAGG 379
Db 3983 GTGGCTACGCGCGGGGCGCGCTACGCTCTCCGTGCAGG 4030

RESULT 6
HSA314900      19475 bp      DNA      linear      PRI 21-MAY-2002
LOCUS      HSA314900      19475 bp      DNA      linear      PRI 21-MAY-2002
DEFINITION Homo sapiens partial OBSCN gene for obscurin, exons all-a16.
ACCESSION   AJ314900
VERSION     AJ314900.1 GI:21104329
KEYWORDS    OBSCN gene; obscurin.
SOURCE      Homo sapiens (human)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS     Young,P., Ehler,E. and Gautel,M.
TITLE       Obscurin, a giant sarcomeric Rho-GEF protein involved in sarcomere
assembly
JOURNAL     Unpublished
REFERENCE 2 (bases 1 to 19475)
AUTHORS     Gautel,M.S.
TITLE       Direct Submission
JOURNAL     Submitted (22-MAY-2001) Gautel M.S., Physikalische Biochemie,
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse
11, Dortmund, 44227, GERMANY
FEATURES
source      1..19475
/organism="Homo sapiens"

```



```
gene      /mol_type="genomic DNA"
          /db_xref="taxon:9606"
          4472..18215
          /gene="OBSCN"
mRNA      join(<4472..4747,5585..5860,6702..6977,7817..8092,
          12544..11529,17940..>18215)
          /gene="OBSCN"
          /gene="OBSCN"
          join(<4472..4747,5585..5860,6702..6977,7817..8092,
          11254..11529,17940..>18215)
          /gene="OBSCN"
          /codon_start=3
          /product="obscurin"
          /protein_id="CAC85747.1"
          /db_xref="GI:21104330"
          /db_xref="UniProt/TreMBL:Q8NHN7"
          /translation="PKVVFADQVAHSEVQAPAGASATLSCEVAQAQTEVVMVYKDGKK
          LSSLLKVVHVEAGRRRLVVOQAGTKTDAGDYSCARGQVSRFLHITGPKMFMFAKEQS
          VNEVQAEAGASAMLSCEVAQAQTEVTWKDGKLSSSKVMGEVQAEAGTTLRLVLPQAG
          KADAGEYSCAGGQRVSRFLHITGPKMFMFAKEQSVMGEVQAEAGTTLRLVLPQAG
          EVTYKDGKLSSSKVMGEVQAEAGTTLRLVVOQAGTKTDAGDYSCARGQVSRFLHITG
          PKVFAKEQLVHNEVTEAGASATLSCEVAQAQTEVTWKDGKLSSSKVMGEVQAEAGT
          NQQLVVOQAGADAGEYTCAGGQRSLRFLDVSEPKVFAKEQLAHRKVMQAEAGAT
          LSCVAQAQTEVTWKDGKLSSSKVMGEVQAEAGTTLRLVVOQAGADAGEYTCAGGQ
          RLSFLDVAEPKVFAKEQPVHREVQAGASATLSCEVAQAQTEVVMVYKDGKLSFS
          SKVRMEVGTTLRLVVOQAGQVAGEYSCAGGSKLSFLHLVA"
          <4472..4747
          /gene="OBSCN"
          /note="a11"
          4748..5584
          /gene="OBSCN"
          /note="a11"
          5585..5860
          /gene="OBSCN"
          /note="a12"
          5861..6701
          /gene="OBSCN"
          /note="a12"
          6702..6977
          /gene="OBSCN"
          /note="a13"
          6978..7816
          /gene="OBSCN"
          /note="a13"
          7817..8092
          /gene="OBSCN"
          /note="a14"
          8093..11253
          /gene="OBSCN"
          /note="a14"
          11254..11529
          /gene="OBSCN"
          /note="a15"
          11530..17939
          /gene="OBSCN"
          /note="a15"
          17940..>18215
          /gene="OBSCN"
          /note="a16"
```

ORIGIN

```
Query Match      22.9%; Score 229.6; DB 9; Length 19475;
Best Local Similarity 84.1%; Pred. No. 4e-29;
Matches 259; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 650 CTTGAGCCCAAGTGTGTTTGTCTAAGGAGCAGCTGGCACGAGGAGCTGCAGGCGAGA 709
      |||
Db 11249 CCCAGAGCCCAAGCGGTGTTTGCACAAAGGAGCAGCTGGCACAGGAGGTGCAGGCCGA 11308
      |||

Qy 710 GGCAGGAGCCAGTCCACACTAGCTCGAGAGTGGCCCGCCAGCCGACAGCGAGTGACGTG 769
      |||
Db 11309 GGCAGGAGCCAGTTCACACGCTGAGCTCGAGAGTGGCCCGCCAGCCGACAGAGTGACGTG 11368
      |||

Qy 770 GTACAGGATGGGAAGAAGCTGAGCTCCAGCTCGAAAGTGTGCATGGAGGCCACAGGCTG 829
```

```
Db 11369 GTACAGGACGGGAAGAGCTGAGCTCCAGTCCAAAGTTCGAATGGAGCGTGTGGGCTG 11428
      |||
Qy 830 CACGCGCAGGCTGTTTGTGACGAGCGAGCGACCGCGGATCCCGGGGAGTATAGTTCGGA 889
      |||
Db 11429 CACACGGAGGCTGCTGTGTCAGCAGGATGATCCAGCGGCGACACCGGGGAGTATAGTTCGA 11488
      |||
Qy 890 GGCTGGGGGCGCAGCGGCTCTCTTCATCTGGATGTCGAAGAGCGCCAGGTTGTTGTC 949
      |||
Db 11489 GGCAGGGGGCGCAGCGGCTCTCTTCAGCTGGACGTCGTCAGTCTTGTGGGCAC 11548
      |||
Qy 950 CAAGGACC 957
      |||
Db 11549 CGAGGAGC 11556
      |||

RESULT 7
AC099089/c
LOCUS      Rattus norvegicus clone CH230-154E3, WORKING DRAFT SEQUENCE.
DEFINITION AC099089.5 GI:30522701
ACCESSION AC099089
VERSION    HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
KEYWORDS   Rattus norvegicus (Norway rat)
SOURCE     Rattus norvegicus
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE 1 (bases 1 to 260998)
AUTHORS   Muzny,D.,Marie., Metzker,M.,Lee., Abramson,S., Adams,C., Alder,J.,
           Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
           Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
           Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
           Biswal,K., Blair,J., Burch,P., Burrell,K., Brown,M.,
           Bryant,N., Buhay,C., Carter,K., Cavazos,I., Ceasar,H., Calderon,E.,
           Cardenas,V., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
           Chacko,J., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
           Cleveland,C., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
           Delgado,O., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
           Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
           Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
           Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
           Georegeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
           Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
           Harvey,X., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
           Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
           Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
           Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
           Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
           Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
           Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
           Lorensuewa,L., Loulseghe,H., Lozano,R.J., Lu,X., Ma,J.,
           Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
           Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
           Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
           Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
           Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Naif,L.,
           Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
           Nwankweme,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
           Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
           Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,
           Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
           Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
           Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
           Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,
           Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,
           Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
           Steelme,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
           Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,
           Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
           Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
```



```

QY 770 GTACAAGATGGGAAGAGCTGAGCTCCAGCTCGAAAGTGTGCATGTGAGGCGCACAGGCTG 829
Db 1011 GTACAAGATGGGAAGAGAGCTGAGTTCCAGCTCGAAAGTGTGCGGTGAGGCGCGTGGCTG 1070

QY 830 CACGCGCAGGCTGTGTGTCAGCAGCAGGCGCAGGCGGATGCCGGGAGTATAGCTGCGA 889
Db 1071 CACACGGAGGCTGTGTGTCAGCAGCAGGCGGCGCAGCAGGCGCGGGAGTACAGCTGCGA 1130

QY 890 GGCTGGGGCCAGCGGCTCTCTTCCATCTTGGATGTCAAAGAGCCCAAGGTGGTGTTCG 949
Db 1131 GGCAGGGGCTAGCAGCTCTCTTCCGCTGCGAGTGGCAGTCTAGTCTTTGGGATGC 1190

QY 950 CAAG 953
Db 1191 TGAG 1194

RESULT 9
AX883644 AX883644 2488 bp DNA linear PAT 17-DEC-2003
LOCUS Sequence 18549 from Patent EP1074617.
DEFINITION AX883644
ACCESSION AX883644
VERSION AX883644.1 GI:40038545
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Ota, T., Isogai, T., Nishikawa, T., Hayashi, K., Saito, K., Yamamoto, J.,
Ishii, S., Sugiyama, T., Wakamatsu, A., Nagai, K. and Otsuki, T.
TITLE Primers for synthesizing full-length cDNA and their use
JOURNAL Patent: EP 1074617-A 18549 07-FEB-2001;
Research Association for Biotechnology (JP)
FEATURES
source 1.2488
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
697..1197
/note="unnamed protein product"
/codon_start=1
/protein_id="CAE92053.1"
/db_xref="GI:40038546"
/translation="MALHSGWRNDCSCAGRWGFSOLHGQCOAWEGSVPTFFSRPSY
PRVFLTWAVSLTSCVPPSPSPKVFPAEQPAHREVQAEAGASATLSCEVAQAOTE
VTWKDQKLLSSSKVRVEAVGCTFRLVVDQAGQAEVSCREAGGQQLSFRLQVAGQ
CFGDAE"

ORIGIN
Query Match 22.4%; Score 224; DB 6; Length 2488;
Best Local Similarity 83.6%; Pred. No. 4.6e-28;
Matches 254; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 650 CTCGTAGCCCAAGTGTGTGTGTATAGGACAGCTGGCAGCAGGAGCTGCAGGCGA 709
Db 891 CCCAGAGCCCAAGTGTGTGTGTGCAAGGAGCAGCCAGCAGCAGGAGGTGCAGGCTGA 950

QY 710 GGCAGGAGCCAGTCCACACTGAGCTGCGAGTGGCGCCAGCCAGCGGAGGTGACGTG 769
Db 951 GGCAGGAGCCAGTCCACACTGAGCTGCGAGTGGCGCCAGCCAGCGGAGGTGACGTG 1010

QY 770 GTACAAGATGGGAAGAGCTGAGCTCCAGCTCGAAAGTGTGCATGTGAGGCGCACAGGCTG 829
Db 1011 GTACAAGATGGGAAGAGAGCTGAGTTCCAGCTCGAAAGTGTGCGGTGAGGCGCGTGGCTG 1070

QY 830 CACGCGCAGGCTGTGTGTCAGCAGGAGCGCAGGCGGATGCCGGGAGTATAGCTGCGA 889
Db 1071 CACACGGAGGCTGTGTGTCAGCAGGAGCGGCGCAGCAGGCGCGGGAGTACAGCTGCGA 1130

QY 890 GGCTGGGGCCAGCGGCTCTCTTCCATCTTGGATGTCAAAGAGCCCAAGGTGGTGTTCG 949
Db 1131 GGCAGGGGCTAGCAGCTCTCTTCCGCTGCGAGTGGCAGTCTAGTCTTTGGGATGC 1190

```

QY 950 CAAG 953
Db 1191 TGAG 1194

RESULT 10
AK024186

LOCUS Homo sapiens cDNA FLJ14124 fis, clone MAMWA1002498.
DEFINITION AK024186
ACCESSION AK024186
VERSION AK024186.1 GI:10436503
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

AUTHORS Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,

Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,

Sekine, M., Oabayashi, M., Nishi, T., Shibahara, T., Tanaka, T.,

Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,

Nagahari, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,

Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H.,

Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,

Kikkawa, E., Omura, Y., Abe, K., Kamiyama, K., Katsuta, N., Sato, K.,

Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H.,

Hirakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M.,

Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,

Yosida, M., Hata, H., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,

Hara, K., Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R.,

Takeuchi, K., Arita, M., Imose, N., Musashino, K., Yuuki, H., Oshima, A.,

Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T.,

Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,

Teraehima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,

Goto, F., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,

Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,

Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,

Tashiro, H., Tanigami, A., Fujimura, T., Ono, T., Yamada, K., Fujii, Y.,

Ozaki, K., Hirao, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,

Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T.,

Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,

Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,

Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,

Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,

Nagase, T., Shimura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,

Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S.

Complete sequencing and characterization of 21,243 full-length

human cDNAs

Nat. Genet. 36 (1), 40-45 (2004)

14702039

REFERENCE 2

AUTHORS Isogai, T., Ota, T., Hayashi, K., Sugiyama, T., Otsuki, T., Suzuki, Y.,

Nishikawa, T., Nagai, K., Sugano, S., Takahashi-Fujii, A., Hara, H.,

Tanase, T., Nomura, Y., Togiya, S., Komai, F., Hara, R., Takeuchi, K.,

Arita, M., Nabekura, T., Ishii, S., Kawai, Y., Saito, K., Yamamoto, J.,

Wakamatsu, A., Nakamura, Y., Nagahari, K., Masuho, Y. and Oshima, A.

NEDO human cDNA sequencing project

Unpublished

3 (bases 1 to 2488)

Isogai, T. and Otsuki, T.

Direct Submission

TITLE Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute,

Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan

(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)

NEDO human cDNA sequencing project supported by Ministry of

International Trade and Industry of Japan; cDNA full insert

sequencing; Research Association for Biotechnology; cDNA library

construction, 5'- & 3'-end one pass sequencing and clone selection;

Helix Research Institute (supported by Japan Key Technology Center

etc.) and Department of Virology, Institute of Medical Science,

University of Tokyo.

LOCATION/Qualifiers

FEATURES


```

source      1. 2488
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="MAMMA1002498"
            /tissue_type="Mammary gland"
            /clone_lib="MAMMA1"
            /note="Cloning vector: pME18SFL3"
            697. .1197
            /codon_start=1
            /protein_id="BAB14847.1"
            /db_xref="GI:10436504"
            /translation="MALHSGRNCDCSAGRHWGFSQLHGQCAWEGSVSPFFSRPSY
            PRVFLWAVSLTSCVPSPKPVFAKEQPAHREVQAGASATLSCVEVAQTE
            VTWYDKGKLSSSKVRVEAVGCTRRLLVVOQAQAEYSCEAGGQGLSFLQVAGQ
            CFGDAE"

CDS
Query Match      22.4%; Score 224; DB 9; Length 2488;
Best Local Similarity 83.6%; Pred. No. 4.6e-28;
Matches 254; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

Qy 650 CTCTGAGCCCAAGGTGGTGTTCCTAAGGAGCAGCTGGCAGCAGAGCTGCAGGCAGA 709
    |||||
Db 891 CCCAGAGCCCAAGGTGGTGTTCCTAAGGAGCAGCTGGCAGCAGAGCTGCAGGCAGA 950
    |||||

Qy 710 GGCAGGAGCCAGTGCACACTGAGCTGCGAGGTGGCCAGCCAGCCAGCGAGGTGACGTG 769
    |||||
Db 951 GCGCGGGCCAGTGCACAGCTGAGCTGCGAGGTGGCCAGCCAGCCAGCGAGGTGACGTG 1010
    |||||

Qy 770 GTACAGAGTGGGAAGAAGCTGAGCTCAGCTCGAAGTGTGCATGTGAGGAGCCACAGGCTG 829
    |||||
Db 1011 GTACAAGAGTGGGAAGAAGCTGAGTTCAGCTCGAAGTGGCGGTGGAGGCGCTGGGGTG 1070
    |||||

Qy 830 CACGCGAGGCTGGTGTGACAGAGGAGCCAGCGAGTCCCGGGAGTATAGCTGCGA 889
    |||||
Db 1071 CACAGGAGGCTGGTGTGACAGAGGAGCCAGCGAGTCCCGGGAGTATAGCTGCGA 1130
    |||||

Qy 890 GGCTGGGGGCGAGCGGCTCTCTTCCTGATCTGATGTCGAAGAGCCCAAGGTGGTGTTCG 949
    |||||
Db 1131 GGCAGGGGCTCAGCAGCTCTCTTCCTGCTGCGAGTGCAGTGCAGTCTCTTGGGGATGC 1190
    |||||

Qy 950 CAAG 953
    ||
Db 1191 TGAG 1194

RESULT 11
AC023889 174612 bp DNA linear HTG 07-JUL-2000
LOCUS Homo sapiens chromosome 1 clone RP11-661B12, WORKING DRAFT
DEFINITION SEQUENCE, 13 unordered pieces.
AC023889
AC023889.3 GI:8969253
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 174612)
AUTHORS Waterston,R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 174612)
AUTHORS Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (18-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Jul 7, 2000 this sequence version replaced gi:8748947.
----- Genome Center -----

```

```

Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0661B12
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-terminator ET; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 163945 bases at least Q40
Consensus quality: 167601 bases at least Q30
Consensus quality: 169687 bases at least Q20
Insert size: 148000; agarose-fp
Insert size: 173412; sum-of-contigs
Quality coverage: 5.67 in Q20 bases; agarose-fp
Quality coverage: 5.55 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2478: contig of 2478 bp in length
* 2479 2578: gap of unknown length
* 2579 6948: contig of 4370 bp in length
* 6949 7048: gap of unknown length
* 7049 9373: contig of 2325 bp in length
* 9374 9473: gap of unknown length
* 9474 13303: contig of 3830 bp in length
* 13304 13403: gap of unknown length
* 13404 18270: contig of 4867 bp in length
* 18271 18370: gap of unknown length
* 18371 25590: contig of 7219 bp in length
* 25590 25690: gap of unknown length
* 25690 34816: contig of 9126 bp in length
* 34816 34915: gap of unknown length
* 34915 34916: contig of 10827 bp in length
* 34916 45743: gap of unknown length
* 45743 45843: contig of 17145 bp in length
* 45843 62988: gap of unknown length
* 62988 63088: contig of 24497 bp in length
* 63088 87584: gap of unknown length
* 87584 87685: contig of 27081 bp in length
* 87685 114766: gap of unknown length
* 114766 114865: gap of unknown length
* 114865 142334: contig of 27469 bp in length
* 142334 142435: gap of unknown length
* 142435 174612: contig of 32178 bp in length.
* 174612 Location/Qualifiers
* 1. 174612
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /chromosome="1"
  /clone="RP11-661B12"
  1. 2478
  /note="assembly_name:Contig4"
  2579. .6948
  /note="assembly_name:Contig5"
  7049. .9373
  /note="assembly_name:Contig6"
  9474. 13303
  /note="assembly_name:Contig7"
  13404. .18270
  /note="assembly_name:Contig8"
  18371. .25589
  /note="assembly_name:Contig9"
  25590. .34815
  /note="assembly_name:Contig10"

```



```

misc_feature 34916..45742
/notes="assembly_name:Contig11"
misc_feature 45843..62987
/notes="assembly_name:Contig12
clone_end:SP6
vector_side:right"
misc_feature 63088..87584
/notes="assembly_name:Contig13"
misc_feature 87685..114765
/notes="assembly_name:Contig14"
misc_feature 114866..142334
/notes="assembly_name:Contig15"
misc_feature 142435..174612
/notes="assembly_name:Contig16"
ORIGIN
Query Match 22.4%; Score 224; DB 2; Length 174612;
Best Local Similarity 83.6%; Pred. No. 2.6e-28;
Matches 254; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 650 CTCTGAGCCCAAGGTGGTGTTCCTAAGGAGCAGCTGGCCAGCAGGAAAGCTGCAGGCAGA 709
DB 95992 CCACAGAGCCCAAGGTGGTGTTCCTAAGGAGCAGCAGCAGCAGGAGGTGCAGGCTGA 96051
QY 710 GCACAGAGCAGTGCACACTGAGCTGCAGCTGCGAGTGCGCCAGGCCAGACGAGGTGACGTG 769
DB 96052 GCGCGGGGCGCAGTGCACACTGAGCTGCAGCTGCAGTGCAGTGCAGTGCAGTGCAGTGC 96111
QY 770 GTACAAGAGTGGGAAGAGCTGAGTCCAGCTCGAAGTGTGCAAGTGGCGGTGGAGGCGGCTG 829
DB 96112 GTACAAGAGTGGGAAGAGCTGAGTCCAGCTCGAAGTGTGCAAGTGGCGGTGGAGGCGGCTG 96171
QY 830 CACGCGCAGGCTGGTGTGTCAGCAGCAGGCGCAGGCGGATGCCGGGGAGTATAGCTGCGA 889
DB 96172 CACACGAGGCTGGTGTGTCAGCAGCAGGCGGCGCAGCAGGCGGAGTACACTGCGA 96231
QY 890 GCCTGGGGCGCAGCGGCTCTCTCCATCTGATGTCAAAGGCCCAAGGTGGTGTTCG 949
DB 96232 GCGAGGGGGTCAAGCAGCTCTCTCCGCTGCGAGTGCAGTGCAGTGCAGTGCAGTGC 96291
QY 950 CAAAG 953
DB 96292 TCAG 96295
RESULT 12
CQ851220/c
LOCUS 3956 bp DNA linear PAT 23-AUG-2004
DEFINITION Sequence 1689 from Patent EP1447413.
ACCESSION CQ851220
VERSION CQ851220.1 GI:51509432
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Isogai,T., Yamamoto,J., Nishikawa,T., Isono,Y., Sugiyama,T.,
Otsuki,T., Wakamatsu,A., Ishii,S., Nagai,K. and Irie,R.
Full-length human cDNA
Patent: EP 1447413-A 1689 18-AUG-2004;
Research Association for Biotechnology (JP)
FEATURES
source
Location/Qualifiers
1..3956
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 22.1%; Score 221.6; DB 6; Length 3956;
Best Local Similarity 86.3%; Pred. No. 1.1e-27;
Matches 245; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

```

```

QY 647 CGTCTCTGAGCCCAAGGTGGTGTTCCTAAGGAGCAGCTGCACGCGAGAGTGCAGGC 706
DB 1135 CATCCAGAGCCCAAGGTGGTGTTCCTAAGGAGCAGCAGCGGAGGTGCAGGC 1076
QY 707 AGAGGCGAGGAGCCAGTGCACACTGAGCTGCAGCTGCAGTGCAGTGCAGTGCAGTGCAG 766
DB 1075 GGAGGCGGGGCGCAGTGCACACTGAGCTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 1016
QY 767 GTGTGTACAGATGGGAAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 826
DB 1015 ATGTTACAAGGACGGGAAGAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAG 956
QY 827 CTGCACGCGCAGGCTGGTGTGTCAGCAGGAGGCGGCGGATGCCGGGAGTATAGCTG 886
DB 955 CTACACACGAGGCTGGTGTGTCAGCAGGAGGCGGCGGATGCCGGGAGTATAGCTG 896
QY 887 CGAGGCTGGGGGCGCAGGCGGCTCTCTTCCATCTGGATGTCAAAG 930
DB 895 TGAGGCGGGGCGCAGGCGGCTCTCTTCCATCTGGATGTCAAAG 952
RESULT 13
AKI28447/c
LOCUS 3956 bp mRNA linear
DEFINITION Homo sapiens cDNA FLJ46590 fis, clone THYMU3044441.
ACCESSION AKI28447
VERSION AKI28447.1 GI:34535823
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
Kodaira,H., Furuya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y.,
Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahara,K., Masuho,Y., Nagai,K. and Isogai,T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 3956)
Isogai,T. and Yamamoto,J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: Reverse Proteomics Research Institute, HRI and
RAB.
FEATURES
source
Location/Qualifiers
1..3956
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="THYMU3044441"
/tissue_type="thymus"
/clone_lib="THYMU3"
/notes="cloning vector: pME18SFL3"
1014..1415
/notes="unnamed protein product"
/codon_start=1
/protein_id="EAC87445.1"
/db_xref="GI:34535824"
CDS

```


/translation="WSPSPGPPHSSAWHPPPPPPPCMPAAPPWOTPPWALCWTRR
DAGSEAEQTPQPMQPSDSDSQAAGLSIIKQCTVTHRAPPTTRAKGLRHVSGTIIISLIG
NRMRPLCLGLGPGWPHLSPALPITQGHRO"

ORIGIN

Query Match 22.1%; Score 221.6; DB 9; Length 3956;
Best Local Similarity 86.3%; Pred. No. 1.1e-27;
Matches 245; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 647 CGTCTCTAGAGCCCAAGGTGGTTTGTCTAAGAGAGAGTGCAGCAGGAAGTGCAGGC 706
|||
Db 1135 CATCCAGAGCCCAAGGTGGTTTGTCCCAAGAGAGCAGCGGCATGCAGCGAGGTGCAGGC 1076
|||
Qy 707 AGAGCGAGGAGCCAGTGCACACTGAGCTGCGAGGTGCCCCAGGCCAGAGCGAGGTGCAC 766
|||
Db 1075 GGAGCGGGGGCCAGTGCACGCTGAGCTGTGAGGTGCCCGCCAGGCGCAGATGGAGGTGCAC 1016
|||
Qy 767 GTGTGTACAAGGATGGGAAGAAGCTGAGCTCCAGCTCGAAAGTGTGTCATGGAGGCCACAGG 826
|||
Db 1015 ATGTGTACAAGGAGCGGAGAGAGCTGAGCTCCAGCTCAAAAGTGGCGATGGAGGCCAGCGG 956
|||
Qy 827 CTGCAGCGCAGGCTGTGTGTGAGCAGCAGCAGCGCAGCGATGCCGGGAGGTATAGCTG 886
|||
Db 955 CTACACAGGAGGCTGTGTGTGAGCAGCAGCGCGGCGAGCGATGCTGGGAGGTACAGCTG 896
|||
Qy 887 CGAGCGTGGGGCCAGCGGCTCTCTTCATCTGGATGTCAAG 930
|||
Db 895 TGAGCGGGGGGCCAAGCGGCTCTCTCCGCTGCAGCTGCAG 852
|||

RESULT 14

AC098133
LOCUS AC098133
DEFINITION Rattus norvegicus clone CH230-162B20, WORKING DRAFT SEQUENCE, 2
unordered pieces.
AC098133
VERSION AC098133.5 GI:30522688
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus

REFERENCE

1 (bases 1 to 244553)
Murny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Albrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Frazer,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorgis,E., Gear,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,I., Haviak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorenshewa,L., Loulseghe,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Mitosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,

Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokemelehen,O., Okwono,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Paaternak,S., Paul,H., Perez,A., Perez,L., Pfamkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,P.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Shen,H.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Soderstrom,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,B., Ugmani,K.,
Valas,R., Vera,V., Villaseana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.

Direct Submission

Unpublished
2 (bases 1 to 244553)
Worley,K.C.

Direct Submission

Submitted (23-OCT-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

Direct Submission

3 (bases 1 to 244553)
Rat Genome Sequencing Consortium.

Direct Submission

Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

COMMENT

The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using ATag
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM

Web site: http://www.hgsc.bcm.tmc.edu/

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: CH230-162B20

Center clone name: CGMF

----- Summary Statistics

Assembly program: Atlas 3.0;

Consensus quality: 221927 bases at least Q40

Consensus quality: 223181 bases at least Q30

Consensus quality: 223817 bases at least Q20

Estimated insert size: 229995; sum-of-contigs estimation

Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank draft data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 2 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.


```

* 1 242725: contig of 242725 bp in length
* 242726 242825: gap of unknown length
* 242826 244553: contig of 1728 bp in length.

```

FEATURES

Location/Qualifiers

```

1..244553
  /organism="Rattus norvegicus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10116"
  /clone="CH230-162B20"
  22770..113965
    /note="clone_boundary"
    clone_end:117
    site:ECORI
  end_sequence:BH266944"
  23334..24184
    /note="clone_boundary"
    clone_end:Sp6
    site:ECORI
  end_sequence:BH266946"

```

misc_feature

```

1..244553
  /organism="Rattus norvegicus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10116"
  /clone="CH230-162B20"
  22770..113965
    /note="clone_boundary"
    clone_end:117
    site:ECORI
  end_sequence:BH266944"
  23334..24184
    /note="clone_boundary"
    clone_end:Sp6
    site:ECORI
  end_sequence:BH266946"

```

misc_feature

```

1..244553
  /organism="Rattus norvegicus"
  /mol_type="genomic DNA"
  /db_xref="taxon:10116"
  /clone="CH230-162B20"
  22770..113965
    /note="clone_boundary"
    clone_end:117
    site:ECORI
  end_sequence:BH266944"
  23334..24184
    /note="clone_boundary"
    clone_end:Sp6
    site:ECORI
  end_sequence:BH266946"

```

ORIGIN

```

Query Match      22.0%; Score 220.2; DB 2; Length 244553;
Best Local Similarity 86.5%; Pred.No.1.1e-27;
Matches 243; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 650 CTCTGAGCCCAAGTGGTGTGTTGCTTAAGGAGCAGCTGGCAGCAGGAGCTGCAGGCAGA 709
Db 995 CCCAGAGCCCAAGTGGTGTGTTGCTTAAGGAGCAGCAGCAGCAGGAGTGAAGGCAGA 1054

QY 710 GGCAGGACCCAGTGCACACTGAGCTGCGAGTGGCCAGCCAGCCAGCGAGGTGACGTG 769
Db 1055 GCGCGGGGCCAGTGCACACTGAGCTGCGAGTGGCCAGCCAGCCAGCGAGTGAACATG 1114

QY 770 GTACAAGATGGGAAGAGCTGAGCTCCAGCTCGAAAGTGTGCATGAGGAGCCACAGGCTG 829
Db 1115 GTTCAAGACCGGAGAGAGCTGAGCTCCAGCTCCAGGTGCGAGTGAAGGCTCGGGCTG 1174

QY 830 CACGCGCAGGTGTTGTGACAGGAGCAGGAGCCAGCGGATCCCGGGAGTATAGCTGCGA 889
Db 1175 CTCCAGGAGGCTGGTGTGTCAGCAGCGGCGGCAAGCGGATCTGGGAGTACAGCTCGA 1234

QY 890 GCGTGGGGCCAGCGGCTCTCTTCCATCTCGATGTCAAG 930
Db 1235 GCGCGGGGACAGAAGGTCTCTTCCGCTCGAGCTCACAG 1275

```

RESULT 15

```

HSA314903 HSA314903 6400 bp DNA linear PRI 21-MAY-2002
LOCUS Homo sapiens partial OBSCN gene for obscurin, exons C-22.
DEFINITION
ACCESSION AJ314903
VERSION AJ314903.1 GI:21104333
KEYWORDS OBSCN gene; obscurin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

```

1
AUTHORS Young,P., Ehler,G. and Gautel,M.
TITLE Obscurin, a giant sarcomeric Rho-GEF protein involved in sarcomere assembly
JOURNAL Unpublished

```

REFERENCE

```

2 (bases 1 to 6400)
AUTHORS Gautel,M.S.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2001) Gautel M.S., Physikalische Biochemie,
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse
11, Dortmund, 44227, GERMANY

```

FEATURES

```

source
1..6400
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"

```

gene

```

2201..6263
  /gene="OBSCN"

```

mRNA

```

join(<2201..2468,3519..3794,4017..4295,4425..4691,
4819..5088,5998..>6263)
  /gene="OBSCN"

```

CDS

```

join(<2201..2468,3519..3794,4017..4295,4425..4691,
4819..5088,5998..>6263)
  /gene="OBSCN"

```

```

/codon_start=1

```

```

/product="obscurin"

```

```

/protein_id="CAC85750.1"

```

```

/db_xref="GI:21104334"

```

```

/db_xref="UniProt/TsEMBL:O8NH5"

```

```

/translation="VVFADQPVHREVOAEGTSTMLSCVAAOCTEVVMYKDGKKLS
SSSKMRVEAVGCTRRLLVVOEAGQADAGEVSCAAGGQRLSPHLHVAEPKVVFAKEQPAC
REVOAEGASATLSCVDAQOMEVTVYDKGLKSSSKVHMEASGTRRLVVQQAQAA
DAGEVSCAAGGQRLSPHLHVAELEPQISERPCREPLVVKHEHDIILTATLTPTSAAT
VTWLKGVETIRRSKRHETASQDTHLTIVHGAQLDLSAIYSCRVAEGQDPPVQVEEV
AAKFCRLLEPVCVELGCTVTLACELSPACAEVVMRCGNTOLRVKRGQPMVAEGPVRSL
TVLGLRAEDAGEVVCESRDDHTSAOLTVSVPVVVKEMSGLSLVVAEGGRTATQCIVLS
PSDVAVVWFEDGALLQPSKFAISQSGASHLTSIDLVLDAQITVBAEGSSAAL
RVBAEPVLFKKLEPQTVERRSSVTLEVLTRAMPPELWTRNATALAPGNVEIHABG
ARHRLVLHNVGFADRGFFGCEPDDKTQAKLTVE"
<2201..2468
  /gene="OBSCN"
  /notes="C"
  2469..3518
    /gene="OBSCN"
  3519..3794
    /gene="OBSCN"
    /number=17
  3795..4016
    /gene="OBSCN"
    /number=18
  4017..4295
    /gene="OBSCN"
    /number=19
  4296..4424
    /gene="OBSCN"
    /number=20
  4425..4691
    /gene="OBSCN"
    /number=20
  4692..4818
    /gene="OBSCN"
    /number=20
  4819..5088
    /gene="OBSCN"
    /number=21
  5089..5997
    /gene="OBSCN"
    /number=21
  5998..6263
    /gene="OBSCN"
    /number=22

```

ORIGIN

```

Query Match      22.0%; Score 220; DB 9; Length 6400;
Best Local Similarity 85.9%; Pred.No.1.9e-27;
Matches 244; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 647 CGTCTCTGAGCCCAAGTGGTGTGTTGCTTAAGGAGCAGCTGGCAGCAGGAGCTGCAGGC 706
Db 3511 CATCCAGAGCCCAAGTGGTGTGTTGCCAAGGAGCAGCCGGCATGCGAGGAGGTGCAGGC 3570

QY 707 AGAGCAGGAGCAGCTGCACACTGAGCTGCAGGTGCGGCCAGGCCACAGCGAGGTGAC 766
Db 3571 GGAGCGGGGGCCAGTGCACCGCTGAGCTGTGAGGTGGCCAGGGCCAGATGGAGGTGAC 3630

QY 767 GTGGTACAGGATGGGAGAGCTGAGCTCCAGCTCGAAAGTGTGCTGAGGCCACAGG 826
Db 3631 ATGGTACAGGACGGGAGAGCTGAGCTCCAGCTCAAAAGTGCACTGAGGCCACGG 3690

```


Qy 827 CTGCACGGCGCAGGCTGGTTGTGCACAGCGCAGGCCAGCGGATGCCGGGGAGTATAGCTG 886
Db 3691 CTACACACGGAGGCTGGTAGTGCACAGCGCGGCCAGCGGATGCTGGGGAGTACAGCTG 3750
Qy 887 CGAGGCTGGGGGCCAGCGGCTCTCCATCTGGATGTCAAAG 930
Db 3751 TGAGCGGGGGGCCAACGGCTGTCTTCGGCTGCACGTGGCAG 3794

Search completed: March 21, 2005, 05:34:58
Job time : 4518.62 secs

**THIRD EDITION
BLANK (USPTO)**


```
* 1225 3032: contig of 1808 bp in length
* 3033 3132: gap of unknown length
* 3133 4493: contig of 1361 bp in length
* 4494 4593: gap of unknown length
* 4594 5860: contig of 1267 bp in length
* 5861 5960: gap of unknown length
* 5961 7670: contig of 1710 bp in length
* 7671 7770: gap of unknown length
* 7771 9669: contig of 1899 bp in length
* 9670 9769: gap of unknown length
* 9770 11715: contig of 1946 bp in length
* 11716 11815: gap of unknown length
* 11816 14243: contig of 2428 bp in length
* 14244 14343: gap of unknown length
* 14344 16887: contig of 2544 bp in length
* 16888 16987: gap of unknown length
* 16988 19247: contig of 2260 bp in length
* 19248 19347: gap of unknown length
* 19348 21375: contig of 2028 bp in length
* 21376 21475: gap of unknown length
* 21476 25025: contig of 3550 bp in length
* 25026 25125: gap of unknown length
* 25126 28051: contig of 2926 bp in length
* 28052 32054: contig of 3903 bp in length
* 32055 32154: gap of unknown length
* 32155 35716: contig of 3562 bp in length
* 35717 35816: gap of unknown length
* 35817 40891: contig of 5075 bp in length
* 40892 40991: gap of unknown length
* 40992 44027: contig of 3036 bp in length
* 44028 44127: gap of unknown length
* 44128 48899: contig of 4772 bp in length
* 48900 48999: gap of unknown length
* 49000 54255: contig of 5256 bp in length
* 54256 54355: gap of unknown length
* 54356 58994: contig of 4639 bp in length
* 58995 59094: gap of unknown length
* 59095 64390: contig of 5296 bp in length
* 64391 64490: gap of unknown length
* 64491 70865: contig of 6375 bp in length
* 70866 70965: gap of unknown length
* 70966 78667: contig of 7702 bp in length
* 78668 78767: gap of unknown length
* 78768 87207: contig of 8440 bp in length
* 87208 87307: gap of unknown length
* 87308 97858: contig of 10551 bp in length
* 97859 97958: gap of unknown length
* 97959 109520: contig of 11562 bp in length
* 109521 109621: gap of unknown length
* 109622 120908: contig of 11288 bp in length
* 120909 121008: gap of unknown length
* 121009 141477: contig of 20469 bp in length
* 141478 141577: gap of unknown length
* 141578 162672: contig of 21095 bp in length
* 162673 162773: gap of unknown length
* 162773 163974: contig of 1202 bp in length
* 163975 164074: gap of unknown length
* 164075 164766: contig of 692 bp in length.
Location/Qualifiers
1..164766
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-245P10"
1..1124
/note="assembly_name:Contig14"
1225..3032
/note="assembly_name:Contig21"
3133..4493
/note="assembly_name:Contig23"
4594..5860
```

FEATURES
source

```
Query Match 59.8%; Score 239.8; DB 2; Length 164766;
Best Local Similarity 99.2%; Pred. No.1e-38;
Matches 241; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAAGACCAAGTGGTGGCGCCAGGGAGGACGTGGAGCTGCGGTGTGAGCTGTACGGGG 60
Db 60103 GAAGACCAAGTGGTGGCGCCAGGGAGGACGTGGAGCTGCGGTGTGAGCTGTACGGGG 60162
Qy 61 GGAAGCCCGTGCCTGGCTGAAGCAGAGGAGGCATCCGCAAGAGCCAGAAGTATGAT 120
Db 60163 GGAAGCCCGTGCCTGGCTGAAGCAGAGGAGGCATCCGCAAGAGCCAGAAGTATGAT 60222
Qy 121 GTGGTCTCGAGGGCAGCATGGCCATGTGCTATCTCCGGGGGCTCGCTCAAGACGCG 180
Db 60223 GTGGTCTCGAGGGCAGCATGGCCATGTGCTATCTCCGGGGGCTCGCTCAAGACGCG 60282
```

ORIGIN

```
misc_feature
5961..7670
/note="assembly_name:Contig24"
misc_feature
7771..9669
/note="assembly_name:Contig25"
misc_feature
9770..11715
/note="assembly_name:Contig26"
misc_feature
11816..14243
/note="assembly_name:Contig27"
misc_feature
14344..16887
/note="assembly_name:Contig28"
misc_feature
16988..19247
/note="assembly_name:Contig29"
misc_feature
19348..21375
/note="assembly_name:Contig30"
misc_feature
21476..25025
/note="assembly_name:Contig31"
misc_feature
25126..28051
/note="assembly_name:Contig32"
misc_feature
28152..32054
/note="assembly_name:Contig33"
misc_feature
32155..35716
/note="assembly_name:Contig34"
misc_feature
35817..40891
/note="assembly_name:Contig35"
misc_feature
40992..44027
/note="assembly_name:Contig36"
misc_feature
44128..48899
/note="assembly_name:Contig37"
misc_feature
49000..54255
/note="assembly_name:Contig38"
misc_feature
54356..58994
/note="assembly_name:Contig39"
misc_feature
59095..64390
/note="assembly_name:Contig40"
misc_feature
64491..70865
/note="assembly_name:Contig41"
misc_feature
70966..78667
/note="assembly_name:Contig42"
misc_feature
78768..87207
/note="assembly_name:Contig43"
misc_feature
87308..97858
/note="assembly_name:Contig44"
misc_feature
97959..109520
/note="assembly_name:Contig45"
misc_feature
109621..120908
/note="assembly_name:Contig46"
misc_feature
121009..141477
/note="assembly_name:Contig47"
misc_feature
141578..162672
/note="assembly_name:Contig48"
misc_feature
162773..163974
/note="assembly_name:Contig49"
misc_feature
164075..164766
/note="assembly_name:Contig50"
```


RESULT 12
AL645854 211829 bp DNA linear ROD 05-APR-2002
LOCUS Mouse DNA sequence from clone RP23-192P17 on chromosome 11,
DEFINITION complete sequence.
ACCESSION AL645854
VERSION AL645854.10 GI:20068636
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
Smith, M.
Direct Submission
Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:19699553.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e. phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: ENBL; Sw:
SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/C_elegans/wormpep/ RP23-192P17 is
from the RPCI-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6.
Location/Qualifiers
1. 211829
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-192P17"
/clone_lib="RPCI-23"

Query Match 35.6%; Score 142.8; DB 10; Length 211829;
Best Local Similarity 74.4%; Pred. No. 3.2e-19;
Matches 180; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 1 GAAGACCACTGGTGGCGCCAGGAGGACGTGGAGCTGCGCTGTGAGCTGTCCAGCGCG 60
DB 57090 GAGGACCACTGGAGACGCTTGGGAGAGTGTAACTGCTGAGCTGTCCAGGCA 57031

QY 61 GGAACCGCGTGCATGCTGTAAGACAGAGGACCATCCGCAAGCCAGCAAGTATGAT 120
DB 57030 GGCACCTCTGTGCGTGGCTGAAAGACGGAAGGCGCATCAGGAAGCCAGCAAGTATGAC 56971

QY 121 GTGCTGTCGAGGCGACGATGCCATGCTGTCATCCGCGGGCTTCGTCAGGACGCG 180
DB 56970 CTGCTCAGTGAAGGACGCGGCTGTGTTGGTTGTCGCAAGCCCTCACTCAAGATTCA 56911

QY 181 GCGGATACCGTGTGAGGTGGAGGCTTCAGAGACGACAGCCAGCTCCATGTGGAAGAA 240
DB 56910 GGGGAATACCTTGGAGACAGAACTTCGAAAGACATGCCCAAGCTCTGTGTGAAGT 56851

QY 241 AA 242
DB 56850 AA 56849

RESULT 13
AC099089/c

LOCUS Rattus norvegicus clone CH230-154E3, WORKING DRAFT SEQUENCE.
DEFINITION AC099089
ACCESSION AC099089
VERSION AC099089.5 GI:30522701
KEYWORDS HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE 1 (bases 1 to 260998)

AUTHORS

Murphy D.Marie, Metzker M.Lee, Abramson S., Adams C., Alder J.,
Allen C., Allen H., Albrooks S., Amin A., Anguiano D.,
Anyalebechi V., Aoyagi A., Ayodeji M., Baca E., Baden H.,
Baldwin D., Bandaranaike D., Barber M., Barnstead M., Benahmed P.,
Biswal K., Blair J., Blankenburg K., Blyth P., Brown M.,
Bryant N., Buhay C., Burch P., Burrell K., Calderon E.,
Cardenas V., Carter K., Cavazos I., Ceasar H., Center A.,
Chacko J., Chavez D., Chen G., Chen R., Chen Y., Chen Z., Chu J.,
Cleaveland C., Cockrell R., Cox C., Coyle M., Cree A., D'Souza L.,
Delgado M.L., Davis C., Davy-Carroll L., De Anda C., Dederich D.,
Devilla O., Denison S., Detamo C., Ding Y., Dinh H., Divya K.,
Draper H., Dugan-Rocha S., Dunn A., Durbin K., Duval B., Eaves K.,
Egan A., Escotto M., Eugene C., Evans C.A., Falls T., Fan G.,
Fernandez S., Finley M., Flagg N., Forbes L., Foster M., Foster P.,
Fraser C.M., Gabisi A., Ganta R., Garcia A., Garner T., Garza M.,
Gebregorgis E., Geer K., Gill R., Grady M., Guerra W., Guvera W.,
Gunaratne P., Haaland W., Hamil C., Hamilton C., Hamilton K.,
Harvey Y., Havlik P., Hawes A., Henderson N., Hernandez J.,
Hernandez R., Hines S., Hladun S.B., Hodgson A., Hognes M.,
Hollins B., Howells S., Hulyk S., Hume J., Idlebird D., Jackson A.,
Jackson L., Jacob L., Jiang H., Johnson B., Johnson R., Jolivet A.,
Karpachy S., Kelly S., Kelly S., Khan Z., King L., Kovar C.,
Kowis C., Kraft C.L., Lebow H., Levan J., Lewis L., Li Z., Liu J.,
Liu J., Liu W., Liu Y., London P., Longacre S., Lopez J.,
Lorensuhewa L., Louised H., Lozano R.J., Lu X., Ma J.,
Maheshwari M., Mahindratne M., Mahmoud M., Malloy K., Mangum A.,
Mangum B., Mapua P., Martin K., Martin R., Martinez E.,
Mawhinney S., McLeod M.P., McNeill T.Z., Meenen E.,
Milosavljevic A., Miner G., Minja E., Montemayor J., Moore S.,
Morgan M., Morris K., Morris S., Munidasa M., Murphy M., Nair L.,
Nankervis C., Neal D., Newton N., Nguyen N., Norris S., Parks K.,
Nwackemehe O., Okwuonu G., Olarnpunsagoon A., Pal S., Parks K.,
Plopper F., Poindexter A., Popovic D., Primus E., Pu L.-L.,
Puzo M., Quiroz J., Rachlin E., Reeves K., Regier M.A., Reigh R.,
Rivers C., Rokey T., Rojas A., Rose M., Rose R., Ruiz S., Riggs P.,
Rivers J., Savary G., Scherer S., Scott G., Shatman S., Shen H.,
Shetty J., Shvartsbeyn A., Sison I., Sitter C.D., Smafs D.,
Snead A., Sodergren E., Song X.-Z., Sorelle R., Sosa J.,
Steimle M., Strong R., Sutton A., Svatek A., Tabor P., Taylor C.,
Taylor T., Thomas N., Thomas S., Tingey A., Trejos Z., Umant K.,
Valas R., Vera V., Villanana D., Waldron L., Walker B., Wang J.,
Wang Q., Wang S., Warren J., Warren R., Wei X., White F.,
Williams G., Willson R., Wleczyk R., Wooden H., Worley K.,
Wright D., Wright J., Wu J., Yakub S., Yen J., Yoon L., Yoon V.,
Yu F., Zhang J., Zhou J., Zhou X., Zhao S., Dunn D., von
Niederhausern A., Weiss R., Smith D.R., Holt R.A., Smith H.O.,
Weinstock G. and Gibbs R.A.

TITLE Direct Submission

JOURNAL

REFERENCE 2 (bases 1 to 260998)

AUTHORS

Worley K.C.

TITLE

JOURNAL

Submitted (09-NOV-2001) Human Genome Sequencing Center, Department


```

Assembly program: Phrap; version 0.990319
Consensus quality: 97841 bases at least Q40
Consensus quality: 98004 bases at least Q30
Consensus quality: 98090 bases at least Q20
Insert size: 85000; agarose-fp
Insert size: 98113; sum-of-contigs
Quality coverage: 11.51x in Q20 bases; agarose-fp
Quality coverage: 9.97x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*
* 1 28467: contig of 28467 bp in length
* 28468 28567: gap of unknown length
* 28568 31135: contig of 2568 bp in length
* 31136 31235: gap of unknown length
* 31236 41487: contig of 10252 bp in length
* 41488 41587: gap of unknown length
* 41588 74365: contig of 32778 bp in length
* 74366 74465: gap of unknown length
* 74466 96541: contig of 22076 bp in length
* 96542 96641: gap of unknown length
* 96642 98613: contig of 1972 bp in length.
-----
Location/Qualifiers
1. 98613
/organism="Takifugu rubripes"
/mol_type="genomic DNA"
/db_xref="taxon:31033"
/clone="221D8"
/clone_lib="Incyte Genomics"
misc_feature
1. 28467
/note="assembly_fragment
vector_side:left"
28568..31135
/note="assembly_fragment"
31236..41487
/note="assembly_fragment"
41588..74365
/note="assembly_fragment"
74466..96541
/note="assembly_fragment"
96642..98613
/note="assembly_fragment
vector_side:right"
ORIGIN
Query Match 17.3%; Score 69.2; DB 2; Length 98613;
Best Local Similarity 59.8%; Fred. No. 0.00024;
Matches 116; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
Qy 12 GGTGGCGCCAGGGAGGACGTGGAGCTGGCTGTGAGCTGTGCACGGCGGGAGAACCCCGT 71
Db 57893 GGCAGTCCGGCGGAGGACGCTGAGCTGTGAGATCACCAGACGGACGTACCAT 57952
Qy 72 GCATGGGTGAAGGACAGAAAGGCCATCCGCAAGAGCCAGAGTATGATGTGCTGCGA 131
Db 57953 TCGGTGGCTGAAATATGCCATTTTCATCCGGCAGAGCCCAAAGTATGAGTGTGTC 58012
Qy 132 GGGCAGATGCGCATGCTGGTCATCCGGGGGCTCGCTCAAGGCGGGGGGAGTACAC 191
Db 58013 GAAGAACAGGCGGAGCTGGTGTATCAGGAACGATCAGGACTCTGGAGATTACTG 58072
Qy 192 GTGTGAGGTGGAGG 205
Db 58073 CTGCGAGGCAGATG 58086

```

```

RESULT 15
AC023889/c
LOCUS
DEFINITION
AC023889 Homo sapiens chromosome 1 clone RP11-661B12, WORKING DRAFT
SEQUENCE, 13 unordered pieces.
ACCESSION
AC023889.3 GI:8969253
VERSION
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 174612)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 174612)
Waterston,R.H.
Direct Submission
Submitted (18-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On Jul 7, 2000 this sequence version replaced gi:8748947.
-----
Center: Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
-----
Center project name: H_NH0661B12
-----
Sequencing vector: M13; 100%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 163945 bases at least Q40
Consensus quality: 167601 bases at least Q30
Consensus quality: 169687 bases at least Q20
Insert size: 148000; agarose-fp
Insert size: 173412; sum-of-contigs
Quality coverage: 5.67 in Q20 bases; agarose-fp
Quality coverage: 5.55 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2478: contig of 2478 bp in length
* 2479 2578: gap of unknown length
* 2579 6948: contig of 4370 bp in length
* 6949 7048: gap of unknown length
* 7049 9373: contig of 2325 bp in length
* 9374 9474: gap of unknown length
* 9474 13303: contig of 3830 bp in length
* 13304 13403: gap of unknown length
* 13404 18270: contig of 4867 bp in length
* 18271 18371: gap of unknown length
* 18371 22589: contig of 7219 bp in length
* 22590 25690: gap of unknown length
* 25690 34815: contig of 9126 bp in length
* 34816 34916: gap of unknown length
* 34916 45742: contig of 10827 bp in length
* 45743 45842: gap of unknown length
* 45843 62987: contig of 17145 bp in length
* 62988 63087: gap of unknown length
* 63088 87584: contig of 24497 bp in length

```


* 87585 87684: Gap of unknown length
* 87685 114765: Contig of 27081 bp in length
* 114766 114865: gap of unknown length
* 114866 142334: Contig of 27469 bp in length
* 142335 142434: gap of unknown length
* 142435 174612: Contig of 32178 bp in length.

FEATURES

source
1. .174612
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-661B12"
1. .2478
/note="assembly_name:Contig4"
misc_feature 2579. .6948
/note="assembly_name:Contig5"
misc_feature 7049. .9373
/note="assembly_name:Contig6"
misc_feature 9474. .13303
/note="assembly_name:Contig7"
misc_feature 13404. .18270
/note="assembly_name:Contig8"
misc_feature 18371. .25589
/note="assembly_name:Contig9"
misc_feature 25690. .34815
/note="assembly_name:Contig10"
misc_feature 34916. .45742
/note="assembly_name:Contig11"
misc_feature 45843. .62987
/note="assembly_name:Contig12
clone_end:SP6
vector_side:right"
63088. .87584
/note="assembly_name:Contig13"
misc_feature 87685. .114765
/note="assembly_name:Contig14"
misc_feature 114866. .142334
/note="assembly_name:Contig15"
142435. .174612
/note="assembly_name:Contig16"

ORIGIN

Query Match 17.2%; Score 69; DB 2; Length 174612;
Best Local Similarity 56.3%; Fred. NO. 0.00024;
Matches 129; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 5 ACCAGTGGGTGGCGCCAGGGGAGGACGTGGAGCTGCGCTGTGAGCTGTCAAGCGCGGGAA 64
Db 155238 ACCAGGAGGTGGATGAGGAGGACCGCACACTTATGCTGCGAGCTGAGCCGGCGGGTG 155179
QY 65 CGCCCGTGCATGGCTGAAGACAGAGGAGGCCATCCCGCAAGAGCCAGAGTATGATGTGG 124
Db 155178 CGAGCGTGGAGTGGCGCAAGGGCTCCCTACAGCTCTTCCCTTGTGCCAAGTACCAGATGG 155119
QY 125 TCTGCGAGGCGACGATGGCCATGCTGTGTCATCCGGGGGCTCGCTCAAGGACGCGGGCG 184
Db 155118 TGCAGATGGTGCAGCTGCAGAGTGTGTTGTTACCGGAGTGGAGCAGGAGATGCGGGTG 155059
QY 185 AGTACACGTGTGAGGTGAGGGCTTCCAAGACACAGCAGCAGCTCCATGT 233
Db 155058 ACTACACGTGTGACACGGGCCACACGACGAGCATGGCCAGCCTCTCTGT 155010

Search completed: March 21, 2005, 05:35:02
Job time : 1812.95 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:24:10 ; Search time 4966.73 Seconds
(without alignments)
10741.310 Million cell updates/sec

Title: US-10-077-130-6_COPY_13600_14700

Perfect score: 1101
Sequence: 1 cacagcagcacactgtgac.....agcagggaagggcccatgttc 1101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba.*
2: gb_hcg.*
3: gb_in.*
4: gb_on.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pi.*
9: gb_pr.*
10: gb_ro.*
11: gb_ets.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1101	100.0	20435	9 HSAJ2535	AJ002535 Homo sapi
2	384	34.9	18760	9 HSAJ4906	AJ314906 Homo sapi
3	384	34.9	135964	9 AL353593	AL353593 Human DNA
C 4	380.8	34.6	174612	2 AC023889	AC023889 Homo sapi
5	350.8	31.9	164766	2 AC026657	AC026657 Homo sapi
C 6	277.8	25.2	260998	2 AC099089	AC099089 Rattus no
C 7	267.6	24.3	211829	10 AL645854	AL645854 Mouse DNA
8	248.4	22.6	18524	6 CQ730656	CQ730656 Sequence
9	200.8	18.2	174612	2 AC023889	AC023889 Homo sapi
C 10	75.4	6.8	168144	2 CR792456	CR792456 Danio rer
C 11	74.8	6.8	125020	9 AF429315	AF429315 Homo sapi
12	71.2	6.5	2000	6 AX655393	AX655393 Sequence
13	58.4	5.3	125020	9 AF429315	AF429315 Homo sapi
14	53.4	4.9	2534	6 AX834822	AX834822 Sequence
15	53.4	4.9	2534	9 AK097489	AK097489 Homo sapi
16	53.4	4.9	2761	9 BC043916	BC043916 Homo sapi
17	53.4	4.9	4789	9 AB046776	AB046776 Homo sapi
18	53.4	4.9	34667	9 HSAJ14905	AJ314905 Homo sapi
19	50.4	4.6	80182	5 AL772356	AL772356 Zebrafish

C 20	50.4	4.6	184266	2 BX571737	BX571737 Danio rer
C 21	50.2	4.6	1005	6 CQ445545	CQ445545 Sequence
C 22	49.4	4.5	2000	6 AX655393	AX655393 Sequence
C 23	49.4	4.5	31828	2 AC151613	AC151613 Emiliana
C 24	49.4	4.5	157754	10 AC132313	AC132313 Mus muscu
C 25	48.6	4.4	1772	9 HUMVBPBA	L05606 Human myosi
26	48.6	4.4	1775	6 CQ721696	CQ721696 Sequence
27	48.6	4.4	1821	9 BC044226	BC044226 Homo sapi
28	47.6	4.3	5390	9 HSAJ14904	AJ314904 Homo sapi
29	47.4	4.3	5992	10 BC058610	BC058610 Mus muscu
30	47.4	4.3	6021	10 AY237726	AY237726 Mus muscu
31	47.2	4.3	197411	10 AC135961	AC135961 Mus muscu
32	47	4.3	152618	10 AC127583	AC127583 Mus muscu
33	46.6	4.2	3613	9 HSM802220	AL137493 Homo sapi
34	46.6	4.2	3951	9 AY245430	AY245430 Homo sapi
35	46	4.2	27060	3 AB055927	AB055927 Procamb
36	46	4.2	91501	2 AC024344	AC024344 Homo sapi
C 37	46	4.2	147499	2 AY178788	AY178788 Canle fam
C 38	46	4.2	277000	1 SC093109	AL939109 Streptomy
C 39	45.6	4.1	648	12 AY657521	AY657521 Synthetic
40	45.6	4.1	1688	8 AK071050	AK071050 Oryza sat
41	45.6	4.1	1791	8 AK105948	AK105948 Oryza sat
C 42	45.6	4.1	2010	6 E01992	E01992 DNA sequenc
C 43	45.6	4.1	13421	1 AE004881	AE004881 Pseudomon
C 44	45.6	4.1	110000	1 AP006618_33	Continuation (34 o
C 45	45.6	4.1	141079	8 AP002868	AP002868 Oryza sat

ALIGNMENTS

RESULT 1	HSAJ2535	Homo sapiens mRNA for obscurin (OBSCN gene).	20435 bp	mRNA	linear	PRI 14-SEP-2001
LOCUS						
DEFINITION						
ACCESSION	AJ002535					
VERSION	AJ002535.1	GI:15026973				
KEYWORDS	OBSCN gene; obscurin.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE						
AUTHORS	Young, P., Ehler, E. and Gautel, M.					
TITLE	Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor protein involved in sarcomere assembly					
JOURNAL	J. Cell Biol. 154 (1), 123-136 (2001)					
MEDLINE	21342081					
PUBMED	11448995					
REFERENCE						
AUTHORS	Gautel, M.S.					
TITLE	Direct Submission					
JOURNAL	Submitted (31-OCT-1997)					
REMARK	European Molecular Biology Laboratory, Meierhofstr. 1, Heidelberg, 69117, GERMANY					
FEATURES						
source						
	Revised by author 20-JUL-2001					
	Location/Qualifiers					
	1. .20435					
	/organism="Homo sapiens"					
	/mol_type="mRNA"					
	/db_xref="taxon:9606"					
	/chromosome="1"					
	/map="1q42.1"					
	/cell_type="cardiomyocyte"					
	/tissue_type="cardiac"					
	/dev_stage="adult"					
gene	1. .20435					
	/gene="OBSCN"					
CDS	72. .19934					
	/gene="OBSCN"					
	/function="titin binding"					
	/codon_start=1					
	/evidence=experimental					

Qy	301	CAGGTGGCAGCTGGTGAAGATGCTCTCTGGAGCTTGAGTGGTGGCTGAGCTGGTGG	360
Db	13971		
Qy	361	GTCATCTGGCACAAGGAATGAGCGCATCAGACCCGGTGGCGGCTTCGAGTGTGCTCC	420
Db	14031		
Qy	421	CAGGTGGCACAAGGAATGAGCGCATCAGACCCGGTGGCGGCTTCGAGTGTGCTCC	480
Db	14091		
Qy	481	CAGGTGGCACAAGGAATGAGCGCATCAGACCCGGTGGCGGCTTCGAGTGTGCTCC	540
Db	14151		
Qy	541	AGCCGAGCCTCTGTGGATGAGCGCCTCAGCGCATCAGACCCGGTGGCGGCTTCGAGTGTGCTCC	600
Db	14211		
Qy	601	GGTGACCTGACCTACTGTGGAGCGCCTGCGTGGGAAACGTGCGATGAGCGGTGAGGCC	660
Db	14271		
Qy	661	ACGCTGGACTCCATTAGCGAGTGCAGGAGGAGCGCGCTCGCAGCGCGCTGCGACAG	720
Db	14331		
Qy	721	GAGCGAGGAGGCTGGACCTGATCTCTGAAGGCTACTCAGCGCGCATGAGCTGGCC	780
Db	14391		
Qy	781	CGCACTGAGATGCTGACCTCTCACACCAAGCTCTGATGATGAGTCCCGGCGAGCGACC	840
Db	14451		
Qy	841	CCTTCCCTGGTCACTACTCAAGAGGCTGGAGCGCAGGACCTCACCACTGGCCAGC	900
Db	14511		
Qy	901	AAGTTGGGGCCCCAGCAGCCCTCTGTGAAGCCACAGCAGCAGCAGCAGCAGCAGCAGC	960
Db	14571		
Qy	961	GCTGTGCCCCCACCCTGGAGACCTGAGCACCAGCAGCAGCAGCAGCAGCAGCAGCAGC	1020
Db	14631		
Qy	1021	AAGCAGCTGTGAGATCCAGCTGCCCTTAAGGCTACAGGTCGAGGAGGAGATGAAG	1080
Db	14691		
Qy	1081	CAGCAGGAAGGCGCCATGTTCT	1101
Db	14751		

RESULT 2	HS314906	18760 bp	DNA	linear	PRI 21-MAY-2002
LOCUS	HS314906				
DEFINITION	Homo sapiens partial OBSCN gene for obscurin, exons 57-68.				
ACCESSION	AJ314906				
VERSION	AJ314906.1	GI:21104339			
KEYWORDS	OBSCN gene; obscurin.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
AUTHORS	Young, P., Ehler, E. and Gautel, M.				
TITLE	Obscurin, a giant sarcomeric Rho-GEF protein involved in sarcomere assembly				
JOURNAL	Unpublished				

REFERENCE	2 (bases 1 to 18760)
AUTHORS	Gautel, M.S.
TITLE	Direct Submission
JOURNAL	Submitted (22-MAY-2001) Gautel M.S., Physikalische Biochemie, Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse 11, Dortmund, 44227, GERMANY
FEATURES	Location/Qualifiers
source	1..18760
gene	/organism="Homo sapiens"
gene	/mol_type="genomic DNA"
gene	/db_xref="taxon:9606"
gene	149..18738
gene	/gene="OBSCN"
gene	join<149..424,904..1191,1332..1613,2324..2694,4784..5654,7668..7046,16300..16377,16635..16763,17023..17220,17952..17984,18255..18353,18520..>18738)
gene	/gene="OBSCN"
gene	join<149..424,904..1191,1332..1613,2324..2694,4784..5654,7668..7046,16300..16377,16635..16763,17023..17220,17952..17984,18255..18353,18520..>18738)
gene	/gene="OBSCN"
gene	/codon_start=3
gene	/product="obscurin"
gene	/protein_id="CAC85753.1"
gene	/db_xref="GI:21104340"
gene	/db_xref="UniProt/TREMBL:Q8NH2"
gene	/translation="WRLEILPLKNAAVRGAQAQCTTLSEAVPVGASWINGAAVQPDDSWTADGSHHALLLRSEVPHHAGEVTPACRDVAVASARLTVLGLPPDEDAEVVARSHVTVLSWAAPMSDGGGCGCYRVEVKEGATGWRLCHELVPGEVVDGLAPG ETVFRVAAGPVGAGRPVHLPTQVRLAEPKVPVPOPSAPESROVAAGEVDSLELEVVAEGEVIWKGMEIRIOPGGRFEVVSQGRQOMLVIKGTAEADQGEYHGLAQGSICPA AATQVALSPASDEAFQPSLPPAAQEGDHLLEALARKRMSREPTLDSISLPE EDGRSRLPQAEAEVAPDLSEGYSTADELARTADGADLSHTSSDDESAGTSLVTLK KAGPFGTSLASKVGAAPAAQKQOQEPAAVPLGLDLSKDLGDFSMKAAVKI QAAQKVKRMEKQOQEPMSHTFGTAEQVGDALRLCEVVASKADVRARWLKDGVE LTRGRHHIDOLGDTCSLLITGLDRADAGCTYQVSNKFGVTHSA CVVVGSGSEEA ESSGGELDDAFRAARLRHLRFLTKSPAESVDELEFLSADGPAEPREPADMQTYRE DEHFCIRFEALTEARQAVTRFQEMFATLGIVEIKLVEQGRPRVEMCISKETPAPVV PPEPLSLLTSDAAVFLTELQNEQVQDGPVDFCVVTGQPMPSVRFVKDGLLEED DAYINEDQGGHQLIITAVPADMGVYRCIAENSMGVSSTKABLRVLDTLTDYDQAA DATESGVFSAQGLSTSRQEGESTTDEQQLPVVEELDLQVAPGTRIAKFLQKV KGYAPRLYFWDGQPLTASAHIRMTDKKIHLTLEIISVTREDSQYAAVYSNMGAA YSARLLVRGDEPEEPASDVHEQLVPPMLERFTPKKVKKGSITFSVKEGVRVP TVHWREAEARGVLWIGDTPGTPTGVASSAQOQSHLVLLDVGQRHQSTYTCIASNAAGQA LCSASLVSG"
exon	<149..424
exon	/gene="OBSCN"
exon	/number=57
intron	425..903
intron	/gene="OBSCN"
intron	/number=57
exon	904..1191
exon	/gene="OBSCN"
exon	/number=58
intron	1192..1331
intron	/gene="OBSCN"
intron	/number=58
exon	1332..1613
exon	/gene="OBSCN"
exon	/number=59
intron	1614..2323
intron	/gene="OBSCN"
intron	/number=59
exon	2324..2694
exon	/gene="OBSCN"
exon	/number=60
intron	2695..4783
intron	/gene="OBSCN"
intron	/number=60
exon	4784..5654
exon	/gene="OBSCN"
exon	/number=61
intron	5655..6767


```
/db_xref="taxon:9606"
/chromosome="1"
/map="q42.1-43"
/clone="RPS-1139B12"
/clone_lib="RPC1-5"
25928..26054
misc_feature
  /note="Sequence from overlapping clone RP11-520H14
(ALJ359510). Assembly confirmed by restriction digest."
ORIGIN
Query Match      34.9%; Score 384; DB 9; Length 135964;
Best Local Similarity 78.3%; Pred. No. 1.9e-61;
Matches 540; Conservative 0; Mismatches 10; Indels 140; Gaps 1;

QY 1 CACAGCACCACACTGTGACACTGTCTGGGAGCTCCCATGAGTATGAGGCGGTGT 60
DB 104829 CGCAGACCCACACTGTGACACTGTCTTTGGGAGCTCCCATGAGTATGAGGCGGTGT 104888
QY 61 CTCTGTGCTACCGGCTGAGGTGAAGAGGGGGCCACAGGCCACTGCGGTGTGCCAC 120
DB 104889 CTCTGTGCTACCGGCTGAGGTGAAGAGGGGGCCACAGGCCACTGCGGTGTGCCAC 104948
QY 121 GAGCTGTGCTGACCCGAGTGTGTGTGATGTCCTGGCCCGGGGAGACCTTACCGC 180
DB 104949 GAGCTGTGCTGACCCGAGTGTGTGTGATGTCCTGGCCCGGGGAGACCTTACCGC 105008
QY 181 TTCCGTGTGGCAGCTGTGGCCCTGTGGTGTGCGGGAACCGGTTCACTGCCCCAGACA 240
DB 105009 TTCCGTGTGGCAGCTGTGGCCCTGTGGTGTGCGGGAACCGGTTCACTGCCCCAGACA 105068
QY 241 GTGGGCTT----- 249
DB 105069 GTGGGCTTGTGAGTCTTCATTGGGGTCTGGGGAGTTACAGCCCCCACAAGCTGA 105128
QY 250 ----- 249
DB 105129 GGAGGGGATCCCCGGAGTGGGGCTGCACCTGGGTGCATCTTGTCTGATGGGCC 105188
QY 250 -----GCAGACCCACCGAAGCTGTGCTCCCGCAGC 280
DB 105189 TGTCTCTGTCGTGGAGCGTGTCTCCCGCAGCAGACCCAGCGCTGTGCTCCCGCAGC 105248
QY 281 CCTACGCCCCGAGAGCGGAGTGGCAGCTGGTGAAGATCTCTCTGGAGCTTGAGG 340
DB 105249 CCTACGCCCCGAGAGCGGAGTGGCAGCTGGTGAAGATCTCTCTGGAGCTTGAGG 105308
QY 341 TGGTGGCTGAGGCTGGTGAAGTCACTTGGCACAAGGGAATGAGGCGCATCAGCCCGGTG 400
DB 105309 TGGTGGCTGAGGCTGGTGAAGTCACTTGGCACAAGGGAATGAGGCGCATCAGCCCGGTG 105368
QY 401 GCGGTTTCAGAGTGTCTCCAGGTCGGCAACAGATGCTGTGATCAAGGGCTTCACGG 460
DB 105369 GCGGTTTCAGAGTGTCTCCAGGTCGGCAACAGATGCTGTGATCAAGGGCTTCACGG 105428
QY 461 CAGAGACCGGGCGAGTACCACTGTGCGCTGGCTCAGGCTCCATCTGCCCTGGGCTG 520
DB 105429 CAGAGACCGGGCGAGTACCACTGTGCGCTGGCTCAGGCTCCATCTGCCCTGGGCTG 105488
QY 521 CCACCTTCCAGTGGCACTTGAGCCCGAGCT 550
DB 105489 CCACCTTCCAGTGGTGGTGTGTGCTTCCCTGTCT 105518

RESULT 4
AC023889/c
LOCUS AC023889
DEFINITION Homo sapiens chromosome 1 clone RP11-661B12, WORKING DRAFT
SEQUENCE, 13 unordered pieces.
ACCESSION AC023889
VERSION AC023889.3 GI:8969253
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 174612)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 174612)
Waterston,R.H.
Direct Submission
Submitted (18-FEB-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Jul 7, 2000 this sequence version replaced gi:8748947.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0661B12
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 163945 bases at least Q40
Consensus quality: 167601 bases at least Q30
Consensus quality: 169687 bases at least Q20
Insert size: 148000; agarose-fp
Insert size: 173412; sum-of-contigs
Quality coverage: 5.67 in Q20 bases; agarose-fp
Quality coverage: 5.55 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 2478: contig of 2478 bp in length
* 2479 2578: gap of unknown length
* 2579 6948: contig of 4370 bp in length
* 6949 7049: gap of unknown length
* 7049 9373: contig of 2325 bp in length
* 9374 13303: contig of 3830 bp in length
* 13304 13403: gap of unknown length
* 13404 18270: contig of 4867 bp in length
* 18271 25589: contig of 7219 bp in length
* 25590 25690: gap of unknown length
* 25690 34815: contig of 9126 bp in length
* 34816 34916: gap of unknown length
* 34916 45742: contig of 10827 bp in length
* 45743 45843: gap of unknown length
* 45843 62987: contig of 17145 bp in length
* 62988 63087: gap of unknown length
* 63088 87584: contig of 24497 bp in length
* 87585 114765: contig of 27081 bp in length
* 114766 114865: gap of unknown length
* 114866 142334: contig of 27469 bp in length
* 142335 142435: gap of unknown length
* 142435 174612: contig of 32178 bp in length.

Location/Qualifiers
1. 174612
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"

misc_feature	/clone="RP11-661B12"	
1. .2478	/note="assembly_name:Contig4"	
misc_feature	2579. .6948	
misc_feature	/note="assembly_name:Contig5"	
7049. .9373	/note="assembly_name:Contig6"	
misc_feature	9474. .13303	
misc_feature	/note="assembly_name:Contig7"	
13404. .18270	/note="assembly_name:Contig8"	
misc_feature	18371. .25589	
misc_feature	/note="assembly_name:Contig9"	
25690. .34815	/note="assembly_name:Contig10"	
misc_feature	34916. .45742	
misc_feature	/note="assembly_name:Contig11"	
45843. .62987	/note="assembly_name:Contig12"	
misc_feature	clone_end:SP6	
63088. .87584	vector_side:right	
misc_feature	/note="assembly_name:Contig13"	
87685. .114765	/note="assembly_name:Contig14"	
misc_feature	114866. .142334	
misc_feature	/note="assembly_name:Contig15"	
142435. .174612	/note="assembly_name:Contig16"	
ORIGIN		
Query Match	34.8%; Score 380.8; DB 2;	
Best Local Similarity	78.0%; Pred. No. 7.1e-61;	
Matches 538; Conservative 0; Mismatches 12; Indels 140; Gaps 1;		
QY	1	CACAGCAGCCACACTGTGACACTGCTTGGGAGCTCCCATGAGTGATGGAGGGGTGGT 60
Db	144677	CGCAGCAGCCACACTGTGACACTGCTTGGGAGCTCCCATGAGTGATGGAGGGGTGGT 144618
QY	61	CTCTGTGGCTACCGGTGGAGTGAGGAGGGGGCCACAGGCGAGTGGCGGCTGTGCCAC 120
Db	144617	CTCTGTGGCTACCGGTGGAGTGAGGAGGGGGCCACAGGCGAGTGGCGGCTGTGCCAC 144558
QY	121	GAGCTGGTCCCTGGACCGGAGTGTGTGTGATGTCCTGGCCCTGGCCCGGGGAGCTACCGC 180
Db	144557	GAGCTGGTCCCTGGACCGGAGTGTGTGTGATGTCCTGGCCCTGGCCCGGGGAGCTACCGC 144498
QY	181	TTCCGTGTGGCAGCTGTGGGCCCTGTGGTGTCTGGGGAACGGTTTCACTGCCCCGACACA 240
Db	144497	TTCCGTGTGGCAGCTGTGGGCCCTGTGGTGTCTGGGGAACGGTTTCACTGCCCCGACACA 144438
QY	241	GTGGGGCTT----- 249
Db	144437	GTGGGGCTTGGTGGTGTCTTCAATTGGGGTCTGGGGATTTCACAGCCCCCACAAGCTGGA 144378
QY	250	----- 249
Db	144377	GGAGGAGGGATCCCGGAGTGGGGCTGCACCCCTGGGTGCAFTTTTGTGTGATGGGGCC 144318
QY	250	-----GAGAGCCACCGAAGCCTGTGCCTCCCGAGC 280
Db	144317	TGTCCTCTGTCGCTGGAGCTGTTCCCGAGCAGACCCAGAGCCTGTGCTCTCCCGAGC 144258
QY	281	CCTCAGCCCTGAGAGCGGAGTGGAGTGTGTGAAGATGTCTCTCTGGAGCTTGAGG 340
Db	144257	CCTCAGCCCTGAGAGCGGAGTGGAGTGTGTGAAGATGTCTCTCTGGAGCTTGAGG 144198
QY	341	TGTTGGCTGAGGCTGGTGCATCTGGCACAAGGAATGGAGCGCATCCAGCCCGTGG 400
Db	144197	TGTTGGCTGAGGCTGGCGAGGTCACTTGGCACAAGGAATGGAGCGCATCCAGCCCGTGG 144138
QY	401	GGCGGTCGAGGTGGTCTCCCGAGTCCGGAACAGATGCTGGTGATCAAGGGCTTCACGG 460


```
* 7771 9669: contig of 1899 bp in length
* 9670 9769: gap of unknown length
* 9770 11715: contig of 1946 bp in length
* 11716 11815: gap of unknown length
* 11816 12433: contig of 2428 bp in length
* 12434 14343: gap of unknown length
* 14344 16887: contig of 2544 bp in length
* 16888 19987: gap of unknown length
* 19988 19247: contig of 2260 bp in length
* 19248 21347: gap of unknown length
* 21348 21375: contig of 2028 bp in length
* 21376 21475: gap of unknown length
* 21476 25025: contig of 3550 bp in length
* 25026 28125: gap of unknown length
* 28126 28151: contig of 2926 bp in length
* 28152 32054: contig of 3903 bp in length
* 32055 32154: gap of unknown length
* 32155 35716: contig of 3562 bp in length
* 35717 35816: gap of unknown length
* 35817 40891: contig of 5075 bp in length
* 40892 40991: gap of unknown length
* 40992 44027: contig of 3036 bp in length
* 44028 44127: gap of unknown length
* 44128 48899: contig of 4772 bp in length
* 48900 48999: gap of unknown length
* 49000 54255: contig of 5256 bp in length
* 54256 54355: gap of unknown length
* 54356 58994: contig of 4639 bp in length
* 58995 59094: gap of unknown length
* 59095 64390: contig of 5296 bp in length
* 64391 64491: gap of unknown length
* 64491 70865: contig of 6375 bp in length
* 70866 70965: gap of unknown length
* 70966 78667: contig of 7702 bp in length
* 78668 78767: gap of unknown length
* 78768 87207: contig of 8440 bp in length
* 87208 87307: gap of unknown length
* 87308 97858: contig of 10551 bp in length
* 97859 97958: gap of unknown length
* 97959 109520: contig of 11562 bp in length
* 109521 109620: gap of unknown length
* 109621 120908: contig of 11288 bp in length
* 120909 121008: gap of unknown length
* 121009 141477: contig of 20469 bp in length
* 141478 141577: gap of unknown length
* 141578 162672: contig of 21095 bp in length
* 162673 162772: gap of unknown length
* 162773 163974: contig of 1202 bp in length
* 163975 164074: gap of unknown length
* 164075 164766: contig of 692 bp in length.
Location/Qualifiers
1..164766
/organism="Homo sapiens"
/mo_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-245P10"
1..1124
/note="assembly_name:Contig14"
misc_feature 1225..3032
/note="assembly_name:Contig21"
misc_feature 3133..4493
/note="assembly_name:Contig23"
misc_feature 4594..5860
/note="assembly_name:Contig24"
misc_feature 5961..7670
/note="assembly_name:Contig25"
misc_feature 7771..9669
/note="assembly_name:Contig26"
misc_feature 9770..11715
/note="assembly_name:Contig27"
misc_feature 11816..14243
```

FEATURES source

```
Query Match 31.9%; Score 350.8; DB 2; Length 164766;
Best Local Similarity 97.7%; Pred. No. 2.5e-55;
Matches 377; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

Qy 532 GTGGACCTGAGCCAGCCTCTGTGATGAGGCCCTCAGCCAGCTTCCCTCCCGAGGCA 591
Db 11299 GTGGACCTGAGCCAGCCTCTGTGATGAGGCCCTCAGCCAGCTTCCCTCCCGAGGCA 11358

Qy 592 GCCCAGGAGGTGACCTGACCTACTGTGGAGGCCCTTGGCTCGGAAACGTGCGATGAGC 651
Db 11359 GCCCAGGAGGTGACCTGACCTACTGTGGAGGCCCTTGGCTCGGAAACGTGCGATGAGC 11418

Qy 652 CGTGAGCCCGCTGGACTCCATTAGGAGGTGCGAGGAGGACGGCCGCTCGCAGGCGC 711
Db 11419 CGTGAGCCCGCTGGACTCCATTAGGAGGTGCGAGGAGGACGGCC-CTCGCAGGCGC 11477

Qy 712 CTGCCACAGGAGGAGGAGGTGCGACCTGATCTCTGAAGGCTACTCCACGCCCGAT 771
Db 11478 CTGCCACAGGAGGAGGAGGTGCGACCTGATCTCTGAAGGCTACTCCACGCCCGAT 11537

Qy 772 GAGCTGGCCCGCTAGGAGGTGCGACCTCTCACACACACAGCTCTGTGATGAGTCCCGG 831
Db 11538 GAGCTGGCCCGCTAGGAGGTGCGACCTCTCACACACACAGCTCTGTGATGAGTCCCGG 11597
```

ORIGIN

```
misc_feature /note="assembly_name:Contig28"
14344..16887
misc_feature /note="assembly_name:Contig29"
16988..19247
misc_feature /note="assembly_name:Contig30"
19348..21375
misc_feature /note="assembly_name:Contig31"
21476..25025
misc_feature /note="assembly_name:Contig32"
25126..28051
misc_feature /note="assembly_name:Contig33"
28152..32054
misc_feature /note="assembly_name:Contig34"
32155..35716
misc_feature /note="assembly_name:Contig35"
35817..40891
misc_feature /note="assembly_name:Contig36"
40992..44027
misc_feature /note="assembly_name:Contig37"
44128..48899
misc_feature /note="assembly_name:Contig38"
49000..54255
misc_feature /note="assembly_name:Contig39"
54356..58994
misc_feature /note="assembly_name:Contig40"
59095..64390
misc_feature /note="assembly_name:Contig41"
64491..70865
misc_feature /note="assembly_name:Contig42"
70966..78667
misc_feature /note="assembly_name:Contig43"
78768..87207
misc_feature /note="assembly_name:Contig44"
87308..97858
misc_feature /note="assembly_name:Contig45"
97959..109520
misc_feature /note="assembly_name:Contig46"
109621..120908
misc_feature /note="assembly_name:Contig47"
121009..141477
misc_feature /note="assembly_name:Contig48"
141578..162672
misc_feature /note="assembly_name:Contig49"
162773..163974
misc_feature /note="assembly_name:Contig9"
164075..164766
misc_feature /note="assembly_name:Contig7"
```


Db	138332	CTCCGAGGCAGCCAGAGGGTGA	CTTGCA	TCTGCTCTTTGGAGGCCCTTGCTCGGAAGC	138273
Qy	641	GTGCGATGAGCCGTGAGCCCA	CGCTGGA	CTCCATTAGCGAGTGCAGAGGAGGACGGCC	700
Db	138272	GTGCGATGAGCCGGGAGCCCA	CGCTGGA	CTCCATCAGTGAGCTGCCGAGGAAGACAGCC	138213
Qy	701	GCTCGAGCGCTGCCACAGAGG	CAGAGGAGTGGCACCTGAT	CTCTCTGAAGGCTACT	760
Db	138212	GTGTGCAGCATCTGGGCA	GAGGCGCAGAGAGGCGGCTCTCTGACCTCTCTGAGGCGTACT	138153	
Qy	761	CCAGGCGCGATGAGCTGGCCCGCA	CTGAGATGTGA	CACTCTCAACACCAAGCTCTGTGATG	820
Db	138152	CCAGGCGCGATGAGCTCGACGCA	CAGAGGAAGCTGACCTCTCA	CACACCAAGCTCTGTGATG	138093
Qy	821	ATGAGTCCCGGGCAGGACCCCTT	CCCTGGTGACCTTCACTCA	AGAAAGGCTCGGAGGCCACG	880
Db	138092	ACGAGTCTCGGGCTGGCACCCCTT	CCCTTAATTACCTACTCA	AGAAAGGCGCGGGGCTCTGTG	138033
Qy	881	GCACCTCACCACTGGCCAGCA	AGGTTGGGGCCCCAGCAGCC	921	
Db	138032	GGATCTCACCCCTTGGCAGCA	AGGTTGACCTCTCTAGTTGGC	137992	

RESULT	7
AL645854/c	
LOCUS	AL645854
DEFINITION	Mouse DNA sequence from clone RP23-192P17 on chromosome 11, complete sequence.
ACCESSION	AL645854
VERSION	AL645854.10
KEYWORDS	GI:20068636
SOURCE	HTG.
ORGANISM	Mus musculus (house mouse)
	Mus musculus
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;

1 Smith, M.
Direct Submission
Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Apr 7, 2002 this sequence version replaced gi:19699553.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em:, EMBL; Sw:.,
SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/projects/c_elegans/wormpep/ **RP23-192P17** is
from the **RP23** Mouse PAC library
constructed by the group of Pieter de Jong.
For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pBACe3.6.
Location/Qualifiers
1. .211829
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-192P17"
/clone_lib="RPC1-23"

Query Match	24.3%;	Score 267.6;	DB 10;	Length 211829;
Best Local Similarity	80.8%;	Pred. No. 5.3e-40;		
Matches 325;	Conservative 0;	Mismatches 74;	Indels 3;	Gaps 1;
Qy	524	CTCTCCAGGTGCACCTGAGCCGAGCCTCTGTGGATGAGGCC---	CTCAGCCCAAGCTTGC	580
Db	23936	CCTCTCAGTGTATGACCTCAGGCTCTGGGATGAGTCCGGC	CACAGCTTACCTGCTC	23877
Qy	581	CCCCGAGGACGCCACGAGAGGGTGACTGCACCTACTGTGGAGGCCCTTGCTCGGAAC	640	
Db	23876	CTCCGAGGACGCCACGAGAGGGTGACTGCATCTGCTTTGGAGGCCCTTCCCGGAAGC	23817	
Qy	641	GTCCGATAGCGTGCAGGCCACGCTGGACTCCATTAGCGAGCTCCAGAGAGGACGCGC	700	
Db	23816	GC CGCATAGTCCGGAGGCCACGCTGGACTCCATCAGTGAAGCTGCTCGAGGACAGAGCC	23757	
Qy	701	GCTCGAGCGCTGCCACAGGAGGCAGAGGAGGTGGCACCTGATCTCTCTGAAGGCTACT	760	
Db	23756	GTGTGCAGCATCTCGCGCAGGAGGCGAAGAGACAGCTCCCGACCTCTCTGAAGGCTACT	23697	
Qy	761	CCACGGCCGATAGCTTGGCCCGCACTGAGATGCTGACCTCTCACACACAGCTCTCATG	820	
Db	23696	CTACAGCGATAGCTTAGCAGCAGAGGAGGCCAGCTCTCTCACACACAGCTCTGACG	23637	
Qy	821	ATGAGTCCGGGCGAGGCACCCCTTCCCTGGTGCACCTACTCTCAAGAAGGCTCGGAGGCCAG	880	
Db	23636	ATGAGTCTCGGCTGGCACCCCTTCCCTAGTTACTACTCTCAGNAGCCGGGGGTCTG	23577	
Qy	881	GCACCTCACCACTGGCCAGCAAGGTTGGGGCCCCCAGCAGCC	922	
Db	23576	GGATCTCACCCCTTGGCCAGCAAGGTGAGCCTCTTAGTTGGCC	23535	

```

RESULT 8
CQ730656
LOCUS
DEFINITION
CQ730656
ACCESSION
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Venter C.J., Adams,M.C., Li,P.W. and Myers,E.W.
TITLE
Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL
Patent: WO 02060579-A 16590 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
Location/Qualifiers
source
1..18524
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

```

	Query Match	22.6%;	Score 248.4;	DB 6;	Length 18524;
	Best Local Similarity	99.6%;	Pred. No. 3.4e-36;		
	Matches 249;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	CA CAG CAG CCA CACT GT G C A C T G T C T T T G G C A G C T C C C A T A G T G A T T G G A G G C G G T G G T	60		
Db	13422	CG CAG CAG CCA CACT GT G C A C T G T C T T T G G C A G C T C C C A T A G T G A T T G G A G G C G G T G G T	13481		
Qy	61	C T C T G T G G C T A C C G G T G G A G G T G A A G A G G G G G C C A C A G G C C A G T G C G G C C T G T G C A C	120		
Db	13482	C T C T G T G G C T A C C G G T G G A G G T G A A G A G G G G G C C A C A G S C C A G T G C G G C C T G T G C C A C	13541		
Qy	121	G A G C T G G T G C C T G G A C C C C A G T G T G T G A T G G C C T T G C C C C C G G G G A G A C C T A C G C G	180		


```
Db      13542 GAGCTGTGCTGGAACCCGAGTGTGTGTGGATGTCCTGCGCCCGGAGACCTACCCG 13601
Qy      181  TTCCGTGTGGCAGCTGTGGGCCCCCTGTGGGTGCTGGGGAACCGGTTCACCTGCCCCCAGACA 240
Db      13602 TTCCGTGTGGCAGCTGTGGGCCCCCTGTGGGTGCTGGGGAACCGGTTCACCTGCCCCCAGACA 13661
Qy      241  GTGCGGCTTG 250
Db      13662 GTGCGGCTTG 13671

RESULT 9
AC023889
LOCUS   Homo sapiens chromosome 1 clone RP11-661B12, WORKING DRAFT
DEFINITION
ACCESSION AC023889
VERSION   AC023889.3 GI:8969253
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE   Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 174612)
AUTHORS  Waterston,R.H.
TITLE    The sequence of Homo sapiens clone
JOURNAL  Unpublished
REFERENCE 2 (bases 1 to 174612)
AUTHORS  Waterston,R.H.
TITLE    Direct Submission
JOURNAL  Submitted (18-FEB-2000) Genome Sequencing Center, Washington
          University School of Medicine, 444 Forest Park Parkway, St. Louis,
          MO 63108, USA
COMMENT  On Jul 7, 2000 this sequence version replaced gi:8748947.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0661B12
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 163945 bases at least Q40
Consensus quality: 167601 bases at least Q30
Consensus quality: 169687 bases at least Q20
Insert size: 148000; agarose-fp
Insert size: 173412; sum-of-contigs
Quality coverage: 5.67 in Q20 bases; agarose-fp
Quality coverage: 5.55 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 13 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 2478: contig of 2478 bp in length
* 2479: gap of unknown length
* 2579: contig of 4370 bp in length
* 7048: gap of unknown length
* 7049: contig of 2325 bp in length
* 9373: gap of unknown length
* 9374: gap of unknown length
* 13303: contig of 3830 bp in length
* 13304: gap of unknown length
* 18270: contig of 4867 bp in length
* 18271: gap of unknown length
```

```

* 18371 25589: contig of 7219 bp in length
* 25590 25689: gap of unknown length
* 25690 34815: contig of 9126 bp in length
* 34816 34915: gap of unknown length
* 34916 45742: contig of 10827 bp in length
* 45743 45842: gap of unknown length
* 45843 62987: contig of 17145 bp in length
* 62988 63087: gap of unknown length
* 63088 87585: contig of 24497 bp in length
* 87585 87684: gap of unknown length
* 87685 114765: contig of 27081 bp in length
* 114766 142334: contig of 27469 bp in length
* 142335 142434: gap of unknown length
* 142435 174612: contig of 32178 bp in length.
FEATURES
          Location/Qualifiers
            source
              1..174612
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="1"
                /clone="RP11-661B12"
            misc_feature
              1..2478
                /note="assembly_name:Contig4"
            misc_feature
              2579..6948
                /note="assembly_name:Contig5"
            misc_feature
              7049..9373
                /note="assembly_name:Contig6"
            misc_feature
              9474..13303
                /note="assembly_name:Contig7"
            misc_feature
              13404..18270
                /note="assembly_name:Contig8"
            misc_feature
              18371..25589
                /note="assembly_name:Contig9"
            misc_feature
              25690..34815
                /note="assembly_name:Contig10"
            misc_feature
              34916..45742
                /note="assembly_name:Contig11"
            misc_feature
              45843..62987
                /note="assembly_name:Contig12"
            misc_feature
              clone_end:SP6
              vector_side:right
            misc_feature
              63088..87584
                /note="assembly_name:Contig13"
            misc_feature
              87685..114765
                /note="assembly_name:Contig14"
            misc_feature
              114866..142334
                /note="assembly_name:Contig15"
            misc_feature
              142435..174612
                /note="assembly_name:Contig16"
ORIGIN
Query Match      18.2%; Score 200.8; DB 2; Length 174612;
Best Local Similarity 96.7%; Pred. No. 1.2e-27;
Matches 205; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy      890  CACTGCGCAGCAAGGTGGGGCCCCCAGCAGCCCCCTCTGTGAAGCCACAGCAGCAGCAGG 949
Db      27074  CTCTGTGCCCCAGAGTTGGGGCCCCCAGCAGCCCCCTCTGTGAAGCCACAGCAGCAGCAGG 27133
Qy      950  AGCCACTGGCTGCTGTGGCGCCGCCACCTGAGGAGCCTGAGCACCACCAAGACCTGGGTGATC 1009
Db      27134  AGCCACTGGCTGCTGTGGCGCCGCCACCTGAGGAGCCTGAGCACCACCAAGACCTGGGTGATC 27193
Qy      1010  CCTCAATGGACAAGCAGCTGTGAAGATCCAGGCTGCTTTAAGGGCTACAGGTCCCGA 1069
Db      27194  CCTCAATGGACAAGCAGCTGTGAAGATCCAGGCTGCTTTAAGGGCTACAGGTCCCGA 27253
Qy      1070  AGGAGATGAAGCAGCAGGAAGGGGCCCATGTGTC 1101
Db      27254  AGGAGATGAAGCAGCAGGAAGGGGCCCATGTGTC 27285
```



```
RESULT 10
CR792456/c
LOCUS
DEFINITION
Danio rerio clone RP71-18A8, *** SEQUENCING IN PROGRESS ***, 6
unordered pieces.
ACCESSION
CR792456
VERSION
CR792456.2 GI:54019787
KEYWORDS
HTG; HTGS_PHASE1.
SOURCE
Danio rerio (zebrafish)
ORGANISM
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 168144)
McLay,K.
Direct Submission
Submitted (07-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Oct 9, 2004 this sequence version replaced gi:52673298.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk
----- Project Information
Center project name: bz18A8
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 166871 bases at least Q40
Consensus quality: 166987 bases at least Q30
Consensus quality: 167320 bases at least Q20
Insert size: 167644; sum-of-contigs
Insert size: 171122; 5.0% error; agarose-fp
Quality coverage: 8.52x in Q20 bases; sum-of-contigs Quality
coverage: 8.70x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 11201: contig of 11201 bp in length
* 11202 11301: gap of 100 bp
* 11302 22330: contig of 11029 bp in length
* 22331 22430: gap of 100 bp
* 22431 93790: contig of 71360 bp in length
* 93791 93890: gap of 100 bp
* 93891 106259: contig of 12369 bp in length
* 106260 106359: gap of 100 bp
* 106360 137391: contig of 31032 bp in length
* 137392 137491: gap of 100 bp
* 137492 168144: contig of 30653 bp in length.
*
* Location/Qualifiers
*     1..168144
*     /organism="Danio rerio"
*     /mol_type="genomic DNA"
*     /db_xref="taxon:7955"
*     /clone_lib="RP71-18A8"
*     /clone_lib="RPC1-71"
*     1..11201
*     /note="assembly fragment:00159
*     fragment_chain:1"
*     11302..22330
*     /note="assembly fragment:00303
*     fragment_chain:1"
*     22431..93790
*     /note="assembly fragment:01403
*     fragment_chain:2"

misc_feature
1..11201
/note="assembly fragment:00049
fragment_chain:2"
106360..137391
/note="assembly fragment:00928
fragment_chain:2"
137492..168144
/note="assembly fragment:00464
fragment_chain:2"
clone_end:SP6
vector_side:right

ORIGIN
Query Match 6.8%; Score 75.4; DB 2; Length 168144;
Best Local Similarity 52.9%; Pred. No. 0.00017;
Matches 282; Conservative 0; Mismatches 236; Indels 15; Gaps 5;
QY 545 CAGCCTCTGTGGATGAGGCCCTCAGCCACCTTGGCCCCCGAGGACGCCCCAGGAGGCTG 604
DB 39878 CACTTCTCAGATGCGGGCAACCAAGCCCAACTGCTCTCTGAGGCTGCTCTGAGGAG 39819
QY 605 ACCTGCACCTACTGTGGGAGGCCCTGGCTCGGAACGTCGCATGAGCGTCGAGCCACGC 664
DB 39818 ATTTACATGCTCTGTGGGAGGCGAGGCCCAAGAGAGAAGAAATGAGTCGAGAACCCAT 39759
QY 665 TGGACTCATTAGGAGCTGCCAGAGGAGGAGCGCGCTCGCAGCGCTCGCAGGAGG 724
DB 39758 TGGATTCAATTTTCAGAGTTGCCAGAGAAGATG---GCAAAAGACAAAATGCAAAAGAAG 39702
QY 725 CAGAGGAGGTGGCACCTGATCTCTCTGAAGGCTACTCCACGCGCATGAGTGGCGCCGA 784
DB 39701 ATGAAAAGAGCCCTGCCAAAGAAATAAAGTGAGACCTGTTGAGCCAGGCCCAAAATCA 39642
QY 785 CTGAGATGCTGAGCTCTCACACACCACTCTGATGATGAGTCCCGGCGCAGCCACCTT 844
DB 39641 CAGCAGAACCAATCTCT---ATACCAGTTCTGATGAGGAATCTCTGAGTGGAAAGTCAA 39585
QY 845 CCCTGGTCACTACTCAAGAGGCTGGGAGCGCCAGGCACTCACCCTGCGCAGCAGG 904
DB 39584 ---TGGTGTCTTACCTTAAGAAAAGCAGCAAGTC---CTCCATGACAGTGGAAAGTCAA 39531
QY 905 TTGGGGCCCCAGAGCCCTCTGTGAAGCCACAGCAGCAGCAGGACCACTGGCTGTG 964
DB 39530 CTGAGACTGTGCTCTTAAGAAATTTGAT---GAGCACTTCAGATGACTGAGCAATCTG 39474
QY 965 TCGCGCCACCACTGGGAGACCTGAGCACCAGAACCTGGGTGATCCCTCAATGGACAAGG 1024
DB 39473 TTACCAAGACATCTGTGGAGCCACTCATGGAAGCTCAGAATGGGATGAGTTGCAAGAGG 39414
QY 1025 CAGCTGTGAAGATCCAGGCTGCTTTAAGGCTCAAGGTCGCGAGAGATG 1077
DB 39413 CTGCTGTCAAGATTCAAGCAGCCTTCAAGGGTTTAAAGCTTAGGAAGACATG 39361

RESULT 11
AF429315/c
LOCUS
DEFINITION
Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION
AF429315
VERSION
AF429315.1 GI:17646244
KEYWORDS
Homo sapiens (human)
SOURCE
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 125020)
Holmes,S.B., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.
A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)
21583737

misc_feature
1..125020
/note="assembly fragment:00159
fragment_chain:1"
11302..22330
/note="assembly fragment:00303
fragment_chain:1"
22431..93790
/note="assembly fragment:01403
fragment_chain:2"
```


PUBMED	11694876	
REFERENCE	2 (bases 1 to 125020)	
AUTHORS	Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.	
TITLE	Direct Submission	
JOURNAL	Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical	
FEATURES	Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA	
source	location/Qualifiers	
	1..125020	
	/organism="Homo sapiens"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:9606"	
	/chromosome="16"	
	/map="16q24.3; between D16S520 and WI-12410"	
	/note="Isolated from a patient with Huntington's	
repeat_region	Disease-Like 2 (HDL2)"	
	complement(35581..35746)	
	/rpt_type=tandem	
	/rpt_unit="ctg"	
gene	complement(<36507..>36887)	
	/gene="JPH3"	
mRNA	/note="synonym: JP3"	
	complement(<36507..>36887)	
	/gene="JPH3"	
	/product="junctophilin 3"	
CDS	complement(<36507..36887)	
	/gene="JPH3"	
	/note="component of the junctional complex between plasma	
	membrane and endoplasmic reticulum"	
	/codon_start=1	
	/product="junctophilin 3"	
	/protein_id="AA140941.1"	
	/db_xref="GI:17646245"	
	/translation="MSSGGGFNFDGSGYCGWEDKAHGHVCTGPKQGEYTGSGS	
	HGFELGVYTPGSNTYQGTWQAKRHGIGLESKGWYKGEWTHFGKRYGVRECA	
	NGAKYEGTWSNGLQDGYGTETYSDG"	
ORIGIN		
Query Match	6.8%; Score 74.8; DB 9; Length 125020;	
Best Local Similarity	11.9%; Pred. No. 0.00024;	
Matches 115; Conservative 428; Mismatches 416; Indels 10; Gaps 4;		
Qy	110 GCGTGTGCCACGAGCTGGTGCCTGCCACCGAGCTGTGTGGTGGTGCCTGGCCCCGGGG 169	
Db	17851 KSKKKKYAYASKGRKKGTTYWRKMSNRKWRWYASWMCAGMAGAACASRSMKCC 17792	
Qy	170 AGACCTACCCCTTCGCTGTGGAGCTGTGGCCCTGTGGTGTCTGGGGAACCGGTTCACC 229	
Db	17791 WGYRKSAGSKSRGGTGYMKKGGKGSMSKWKGSSTSRRRSSAKSCSYMWGSM 17732	
Qy	230 TGCCCCAGACAGTGGCGCTTGACAGACCACCGAAGCTGTGCTCCCCAGCCTCAGCC 289	
Db	17731 CMSCMSMKSMMYCYCMVRMSMYYYKYVSCMGMSSTSYSYSCWMSWGCSCYKCM 17672	
Qy	290 CTGAGACCGCGCAGGTGGCAGCTGGTGAAGATGCTCTCTGAGAGCTTGAGGTGGTGGCTG 349	
Db	17671 KYISGWSSYSTSMGYYSSTSKKYKCSWSMSYKCKTSYRKRYSYIYWGGRKAKKY 17612	
Qy	350 AGCTGTGAGGTATCTGGCACAAGGAATGAGCGCATCAGCCCGGTGGCGGTTCCG 409	
Db	17611 CAGRRRMSYWKCAKWMMSYCYCW--SYCMTYYYSKSTYKSTCYKRGYYWGSKT CVS 17554	
Qy	410 AGTGTGCTCCACGGGTCCGCAACAGATCTGGTGTATCAAGGCC-----TTACGGCAG 463	
Db	17553 AGGKSRMYIYCMWRSSKSSWSMSWARSWCMGWGAGYRRSRKSGAGWRRSRKGRST 17494	
Qy	464 AAGACGAGCGGAGTACCACTGTGGCCTGGCTCAGGGTCTCATCTGCCTCGCGCTGCCA 523	
Db	17493 GMKRACSKKTGSVSTGRSMKKKGYSKYSRGMKGGKKTCTCYMWKYKYKRTSMCWYYM 17434	
Qy	524 CCTTCCAGGTGGCACTGAGCCAGCCTCTGTGGATGAGCCCTCAGCCCAAGTTGCCCC 583	
Db	17433 KSWGYRYKRCMKKKKGTGYRGMSSKKSGYKSMRGMGSSYSTSCWKSCHGYSWMKCMK 17374	


```

QY 448 AAGCGCTTCACGGCAGAGACAGCGGCGAGTACCACTGTGGCTGGCTCAGGCTCCATC 507
Db 198 SSCGRSGCGRRSAVSRYGTGRKGYTKMTYYSASRCWRAJYMTTSSYAWCSSYTWCRSK 257
QY 508 TGCCCTCGGGCTGCCACCTTCAGGTGGCAGCTGAGCCAGCGCTCT--GTGATGAGGCC 564
Db 258 RRSMMWKKMRKMRWSRSYGNYSWSYKMMCTAYKSYSRWCYMYRGCGWRGATRYWGRG 317
QY 565 CCTCAGCCAGCTTCCCCCGGAGCGACGCCAGGAGGTGACCTGACACTACTGTGGGAG 624
Db 318 YMSRMAMMYKKMYRYGYKGMKRGWAGRMWMSRCRWSKACYMYRWRWMTTRRRRWAK 377
QY 625 GCCCTGGCTCGGAACG---TCGCATGAGCGGTGAGCCACGCTGGACTCCATTAGCCAG 681
Db 378 KSSRTSRKRRKMRKMRKRYKMRGYSRMRSCRRARWKRRCRSGRAWKXGRCGTCRMRK 437
QY 682 CTGCAGAGGAGGAGCGCGCTCGCAGCGCTGCCACAGGAGGAGAGGAGGTGGCAGCT 741
Db 438 SYGMRWKSWMASKYKMSRMYRWRKKCSRTTWGKTGGMGTGRCRYKRSKWK 497
QY 742 GATCTCTGAGGCTACTCCAGCGCCGATGAGCTGGCCGCGCACTGGAGATGCTGACCTC 801
Db 498 RKCRRRRGRMYRMRKRYMSARYTMYRCARKKYSYSAARKARCYRGKGYWAGMMWK 557
QY 802 TCACACACGCTCTGATGATGAGTCCCGGCGAGCGACCCCTTCCCTGGTACCTACCTC 861
Db 558 RYKRMYYMMWYKRYKYSKCSWYKMSYASCMKSARKAGAKMKCRSKMSAWSKMSRSS 617
QY 862 AAGAGGCTGGGAGCGGAGGACCTCACCACTGGCCAGCAAGTGTGGGCGCCAGCAGCC 921
Db 618 RKCRKASRSARKYAMGGTSGRMSRWKSYTCYWRKWSKMTCTTMYMYSKYTYA 677
QY 922 CCCTCTGTGAAGCCACAGCAGCAGGAGGACCACTGGCTGTGTGCGGCCACCACTGGGA 981
Db 678 KYGVSRYRYRAWCMYMRWYRYRYRYMYTMYMAWYTSSTRMAWTGMKYSGRYWTSMYKYC 737
QY 982 GACCTGACACCAAGACCTGGGTGATCCCTCAATGGACA 1021
Db 738 KCSWKYRSMYMYWYWWAKTWKMRRYRATRMWMMYRYSM 777

```

RESULT 13

```

AF429315 AF429315 125020 bp DNA linear PRI 18-JAN-2002
DEFINITION Homo sapiens junctophilin 3 (JPH3) gene, partial cds.
ACCESSION AF429315
VERSION AF429315.1 GI:17646244
KEYWORDS
SOURCE Homo sapiens (human)

```

ORGANISM

```

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

```

1 (bases 1 to 125020)
Holmes,S.E., O'Hearn,E., Rosenblatt,A., Callahan,C., Hwang,H.S.,
Ingersoll-Ashworth,R.G., Fleisher,A., Stevanin,G., Brice,A.,
Potter,N.T., Ross,C.A. and Margolis,R.L.

```

AUTHORS

```

A repeat expansion in the gene encoding junctophilin-3 is
associated with Huntington disease-like 2
Nat. Genet. 29 (4), 377-378 (2001)

```

TITLE

```

associated with Huntington disease-like 2

```

JOURNAL

```

MEDLINE
PUBMED
2 (bases 1 to 125020)
Holmes,S.E., Ingersoll-Ashworth,R.G., Ross,C.A. and Margolis,R.L.

```

AUTHORS

```

Direct Submission
Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical
Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA

```

JOURNAL

```

Location/Qualifiers
1.125020

```

FEATURES

```

source
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="16"

```

```

/map="16q24.3; between D16S520 and WI-12410"
/note="isolated from a patient with Huntington's
Disease-Like 2 (HDL2)"
complement(35581..35746)
/rpt_type=tandem
/rpt_unit="ctg"
complement(<36507..>36887)
/gene="JPH3"
/note="synonym: JP3"
complement(<36507..>36887)
/gene="JPH3"
/product="junctophilin 3"
complement(<36507..>36887)
/gene="JPH3"
/note="component of the junctional complex between plasma
membrane and endoplasmic reticulum"
/codon_start=1
/product="junctophilin 3"
/protein_id="AAL40941.1"
/db_xref="GI:17646245"
/translation="MSSGGRPNFDDGSGYCGGWEDGKAHGHVCTGPKGQGEYTGSGS
HGFEVLGVYTPSGNTYQGTWAQQRHIGLESKRWYKGEWTHGPKRGYGVRECA
NGAKYEGTWSNGLQDGYGTETYSDG"
ORIGIN
Query Match 5.3%; Score 58.4; DB 9; Length 125020;
Best Local Similarity 10.6%; Pred. No. 0.25;
Matches 103; Conservative 406; Mismatches 454; Indels 8; Gaps 1;
QY 1 CACAGCAGCACACTGTGACACTGTCTTGGGCAGCTCCCATGATGATGAGGCGGTGT 60
Db 16998 BMSVSYVMSHSMASBSCHWBKMTWSCNMMYKSSWSSWGMCCNRRRSKGMKW 16957
QY 61 CTCTGTGCTACCGCTGAGGTGAAGAGAGGGGCCACAGCCAGTGGCGCTGTGCCAC 120
Db 16958 YSRGMSRSKRYTGGSKMRSMMCTSSCVASCMCCWCCSCCWCRCSCCMMR 17017
QY 121 GAGTGTGCTGCGACCGAGTGTGTGTGATGCGCTGGCCCGCGGGAGACCTACGCG 180
Db 17018 YCCACKCYMSYWTMSASYSMSRRSYWKRSMKCMWRCMSRSRSCSKSGCGSGMKGKG 17077
QY 181 TTCGTGTGGCAGCTGTGGGCGCTGTGGTCTGGGAACC-----GGTTCACTGC 232
Db 17078 KSYGRKTKRSRGMKAKWYMYRRMRMRKMYRSMYKMYCNYCWCWRRRCYCSCWTSR 17137
QY 233 CCCAGACAGTGGCTTGCAGAGCCACGAAAGCTGTGCTCCCGAGCCCTCAGCCCTG 292
Db 17138 SAMCCSYCAKCKSMCYSCYVGMKGYVYACSYRGSMSKSKYCMRGSTYSTSCGCCCTT 17197
QY 293 AGACCGGCGAGTGGCAGCTGTGGAAGATGTCTCTGGAGCTTGAAGTGGTGGCTGAGG 352
Db 17198 TTCCCCCNANTGGGAAGCTTTTNCNKITYSYRKRNGSCAMCKYNNYNNSSCRAGM 17257
QY 353 CTGTGTAGGTTCATCTGGCACAAGGAATGAGGCGCATCCAGCCGGTGGCGGTTCGAGG 412
Db 17258 SCTYKSGMTMSMAYSCWMSMYCNSMRMSASRGMSWSSYMMKMSWRMSYMKCCWCM 17317
QY 413 TGGTCTCCAGGGTCGGCAACAGATGTGTGTGATCAAGGGCTTCACGCGAGAAGACAGG 472
Db 17318 KMVCMMRSMRSGMSYMYVYASWKSRSRGCTCYCWCMSKSKCYKSYMMWMSRMRMKG 17377
QY 473 GCGAGTACCACTGTGGCTGGCTCAGGCTCCATCTGCGCTGGCGCTGCCACCTTCCAGG 532
Db 17378 KWSRCWMSWMSASRSCKYKSMRCSMMSSKCYRCAGCMWRRGMYRMRCSMKRWR 17437
QY 533 TGGCACTGAGCCAGCTCTGTGTGATGAGGCCCTCAGCCAGCTTGGCCCCCGAGGAG 592
Db 17438 WGSAMYMRMRWKGRRGAMCMCKYSRMSRSMKMSYCASCRSCAMSGTMYMKCASYM 17497
QY 593 CCCAGAGGGTGAACCTGACACCTACTGTGGGAGGCGCTGGCTCGGAAACGTCGCATGAGCC 652
Db 17498 CMSSYCTCWTCTSYMSYRCTCWCCKGWSSTYKSKSWSSMSYKGRKRSYSCMCTSRGA 17557

```



```
QY 653 GTGAGCCACGCTGGACTCATTAGCAGCTGCCAGAGGAGCGCGCTCGCAGCGC 712
Db 17558 MSCWRRCCYVRGASMSRRRAKGRSGWGRSKWMTTGGWRSKYYYCTGRMMWT 17617
QY 713 TGCACAGGAGGACAGAGGTGGCACCTGATCTCTCTGAAGGCTACTCACGCGCGATG 772
Db 17618 YMCCWRRSYYMYRGSANVMGRKSSWGNRMGMSASRRCKRSASRSSWCSRMMKMGRRS 17677
QY 773 AGCTGGCCCGCAGCTGGAGATGCTGACTCTTCACACACCGAGCTGTGATGATGACTCCCGG 832
Db 17678 CWSKMWGGRSRSASCKGSGRMRRRSKYRRRGKRSMTKSGKSGKSKWK 17737
QY 833 CAGCACCCCTTCCCTGGTCACTACTCAAGAGGCTGGAGCCGAGCAGCAGCTCACCAC 892
Db 17738 RSGSMTSSCYYSASSCMMMSKSCWCCMMKRCACCYSSMCTSNYRCOWGMSKY 17797
QY 893 TGGCCAGCAGGTGGGGCCCCAGCAGCCCCCTCTGTGAAGCCACAGCAGCAGGAGGAC 952
Db 17798 STCGTCKKCTGYKRSRTWMMYSWTSKYMMWRAACMMYCYMSTRMMCMMSMCAKGW 17857
QY 953 CACTGGCTGCT 963
Db 17858 CMCTMGAGAT 17868

RESULT 14
AX834822
LOCUS 2534 bp DNA linear PAT 15-DEC-2003
DEFINITION Sequence 1946 from Patent EP1347046.
ACCESSION AX834822
VERSION AX834822.1 GI:39920957
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
AUTHORS
Isogai,T., Sugiyama,T., Otsuki,T., Tanase,T., Imose,N., Tanase,T., Takeuchi,K.,
Yamamoto,J., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R.,
Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahara,K. and
Masuho,Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1347046-A 1946 24-SEP-2003;
RESEARCH Association for Biotechnology (JP)
FEATURES
source
1..2534
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Query Match 4.9%; Score 53.4; DB 6; Length 2534;
Best Local Similarity 53.6%; Pred. No. 5.5;
Matches 111; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
QY 287 CCCCTGAGCGCGCAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 346
Db 803 CCCCTGGAAGACGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 862
QY 347 CTGAGGCTGGTCAAGTCAATCTGGCACAAGGGAATGGAGCGCATCCAGCCGCTGGCGGT 406
Db 863 GGGCGGGAACCCGCTGCACCTGGTGAAGGACAGAGAGGCGCATCCGCAAGCGCAGAGT 922
QY 407 TCGAGGTGGTCTCCAGGTCGGCACAAGATGTGGTGATCAAGGGTTTCAAGGCAAG 466
Db 923 ATGATGTGGTCTGGAGGCGGACGATGGCCATGTGGTCAATCCGCGGGGCTCGCTCAAGG 982
QY 467 ACCAGGCGGAGTACACTGTGGCTCGG 493
Db 983 ACGGCGGCGAGTACACTGTGGAGTGG 1009

RESULT 15
```

AK097489

LOCUS Homo sapiens cDNA FLJ40170 fis, clone TEST12016848, weakly similar to H.sapiens mRNA for titin protein. 2534 bp mRNA linear PRI 30-JAN-2004

AK097489

ACCESSION AK097489.1 GI:21757270

VERSION

KEYWORDS oligo capping; fis (full insert sequence).

KEYWORDS

SOURCE Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Obayashi, M., Nishi, T., Shibahara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yosida, M., Hotuta, T., Kusano, J., Kanehori, K., Takahashi-Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togiya, S., Konai, P., Hara, R., Takeuchi, K., Arita, M., Imose, N., Musashino, K., Youki, H., Oshima, A., Sasaki, N., Aotsuka, S., Yoshioka, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y., Otsuki, K., Hiroa, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, K., Noguchi, S., Itoh, T., Shigetani, K., Senba, T., Matsumura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togoashi, T., Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isogai, T. and Sugano, S. Complete sequencing and characterization of 21,243 full-length human cDNAs

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Musashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

NEDO human cDNA sequencing project
Unpublished
3 (bases 1 to 2534)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUN-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 01-438-52-3975, Fax: 01-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers

source

1..2534
/organism="Homo sapiens"

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:32:32 ; Search time 251.197 Seconds
(without alignments)
9503.485 Million cell updates/sec

Title: US-10-077-130-6_COPY_8500_8900
Perfect score: 401
Sequence: 1 gaagaccaggggtggggcc.....caccagccagcaggagg 401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	Query Match	Length	ID	Description	
1	401	100.0	2534	17	US-10-108-260A-1946	Sequence 1946,
2	401	100.0	7564	18	US-10-476-397-15	Sequence 15, Ap
3	401	100.0	14061	17	US-10-093-463-73	Sequence 73, Ap
4	401	100.0	14109	17	US-10-093-463-71	Sequence 71, Ap
5	401	100.0	20489	18	US-10-723-860-5701	Sequence 5701,
6	401	100.0	23907	13	US-10-077-130-6	Sequence 6, App
7	401	100.0	24120	13	US-10-077-130-4	Sequence 4, App
8	60.6	15.1	20565	10	US-09-764-891-6045	Sequence 6045,
c 9	60.6	15.1	20565	14	US-10-091-438-270	Sequence 270, A
10	59.8	14.9	81940	9	US-09-759-508B-1	Sequence 1, Appl
11	59.8	14.9	81940	10	US-09-960-706-1092	Sequence 1092,

12	59.8	14.9	81940	10	US-09-873-319-724	Sequence 724, App
13	59.8	14.9	81940	18	US-10-723-860-132	Sequence 132, App
14	59.8	14.9	81940	18	US-10-656-873A-1	Sequence 1, Appli
15	59.8	14.9	82027	19	US-10-278-698-1034	Sequence 1034, Ap
16	59.8	14.9	82027	19	US-10-278-698-1045	Sequence 1045, Ap
17	56	14.0	716	10	US-09-822-846-117	Sequence 117, App
18	56	14.0	1266	10	US-09-822-846-116	Sequence 116, App
19	56	14.0	1645	9	US-09-726-643-22	Sequence 22, Appl
20	56	14.0	1645	13	US-10-042-141-22	Sequence 22, Appl
21	56	14.0	1645	19	US-10-919-272-22	Sequence 22, Appl
22	56	14.0	2170	17	US-10-264-049-850	Sequence 850, App
23	56	14.0	2170	10	US-09-764-891-6043	Sequence 6043, Ap
24	56	14.0	9591	14	US-10-091-438-268	Sequence 268, App
25	56	14.0	12415	10	US-09-764-891-6044	Sequence 6044, Ap
26	56	14.0	12415	14	US-10-091-438-269	Sequence 269, App
c 27	55.8	13.9	466	16	US-10-029-386-22631	Sequence 22631, A
28	54.4	13.6	2768	17	US-10-120-988-426	Sequence 426, App
29	47.8	11.9	3935	17	US-10-108-260A-534	Sequence 534, App
30	47.8	11.9	4200	17	US-10-341-434-140	Sequence 140, App
31	47.2	11.8	607	13	US-10-027-632-263233	Sequence 263233,
32	47.2	11.8	607	17	US-10-027-632-263233	Sequence 263233,
c 33	45.6	11.4	2090	18	US-10-437-963-91821	Sequence 91821, A
34	45.6	11.4	2254	9	US-09-726-643-44	Sequence 44, Appl
35	45.6	11.4	2254	13	US-10-042-141-44	Sequence 44, Appl
36	45.6	11.4	2254	19	US-10-919-272-44	Sequence 44, Appl
37	45.2	11.3	471	9	US-09-864-761-2416	Sequence 2416, Ap
38	45.2	11.3	573	9	US-09-864-761-14944	Sequence 14944, A
c 39	44.8	11.2	393	15	US-10-259-165-453	Sequence 453, App
c 40	44.8	11.2	396	15	US-10-259-165-121	Sequence 121, App
41	44.8	11.2	525	18	US-10-437-963-37029	Sequence 37029, A
42	44.8	11.2	2031	15	US-10-156-761-1451	Sequence 2451, Ap
c 43	44.8	11.2	9025608	15	US-10-156-761-1	Sequence 1, Appli
44	44.4	11.1	400	17	US-10-242-535A-8373	Sequence 8373, Ap
45	44.4	11.1	400	17	US-10-085-783A-8373	Sequence 8373, Ap

ALIGNMENTS

RESULT 1

US-10-108-260A-1946
; Sequence 1946, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1946
; LENGTH: 2534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1946

Query Match 100.0%; Score 401; DB 17; Length 2534;
Best Local Similarity 100.0%; Pred. No. 9.7e-109;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAAGACCAAGTGGTGGCGCCAGAGGACGCTGGAGCTGCGGTGTAGCTGTACCGGGCG	60
Db	808	GAAGACCAAGTGGTGGCGCCAGAGGACGCTGGAGCTGCGGTGTAGCTGTACCGGGCG	867
Qy	61	GGACGCCCGTGCACCTGGCTGAAGCAGAGGAGCCATCCGCAAGCCAGCAAGTATGAT	120
Db	868	GGACGCCCGTGCACCTGGCTGAAGCAGAGGAGCCATCCGCAAGCCAGCAAGTATGAT	927
Qy	121	GTGGTCTCGAGGGCAGCATGCTGCTCATCCCGGGGCGCTCGCTCAAGGACGCG	180
Db	928	GTGGTCTCGAGGGCAGCATGCTGCTCATCCCGGGGCGCTCGCTCAAGGACGCG	987

QY 181 GCGAGTACACGTGTGAGGTGGAGGCTTCCAAAGAGCACAGCCAGGCTCCATGTGGAAGAA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
988 GCGGAGTACACGTGTGAGGTGGAGGCTTCCAAAGAGCACAGCCAGGCTCCATGTGGAAGAA 1047
QY 241 AAAGCAAACTGTTTACAGAGGAGTGAACCAATCTGCAAGTGGAGGAGAAAGGCGACAGCT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1048 AAAGCAAACTGTTTACAGAGGAGTGAACCAATCTGCAAGTGGAGGAGAAAGGCGACAGCT 1107
QY 301 GTGTTTCACTGCAAGACGGAGCACCCCGGGGCCACAGTGACCTGGCGCAAGGGGCTCTTTG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1108 GTGTTTCACTGCAAGACGGAGCACCCCGGGGCCACAGTGACCTGGCGCAAGGGGCTCTTTG 1167
QY 361 GAGTACGGGCTCTCAGGGAAGCACAGGCCAGCCAGGAGGG 401
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
1168 GAGTACGGGCTCTCAGGGAAGCACAGGCCAGCCAGGAGGG 1208

RESULT 2

US-10-476-397-15
; Sequence 15, Application US/10476397
; Publication No. US20040115687A1
; GENERAL INFORMATION:

; APPLICANT: YUE, Henry
; APPLICANT: LEE, Ernestine A.
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: THANGAVELOU, Kavitha
; APPLICANT: HONCHELL, Cynthia D.
; APPLICANT: DING, Li
; APPLICANT: JACKSON, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: KALLICK, Deborah A.
; APPLICANT: LEE, Sally
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: TRAN, Uyen K.
; APPLICANT: LAL, Preeti G.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: YAO, Monique G.
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: GANDHI, Aneena R.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Raninder K.
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: CHINN, Anna M.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: RAMKUMAR, Javalaxmi
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: FORSYTHE, Ian J.
; TITLE OF INVENTION: CELL ADHESION AND EXTRACELLULAR MATRIX PROTEINS

; FILE REFERENCE: PF-0968 USN
; CURRENT APPLICATION NUMBER: US/10/476,397
; CURRENT FILING DATE: 2003-10-31
; PRIORITY APPLICATION NUMBER: PCT/US02/13874
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/288,290
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/292,468
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/298,616
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/301,672
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/345,008
; PRIOR FILING DATE: 2002-01-04
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 7564
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

;

OTHER INFORMATION: Incyte ID No: 7326129CB1

US-10-476-397-15

Query Match 100.0%; Score 401; DB 18; Length 7564;
Best Local Similarity 100.0%; Pred. No. 1.1e-108;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAGACCAAGTGGGTGGCGCCAGGGAGGACGTGGAGCTGGCGCTGTGAGCTTCTACGGGCG 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2617 GAAGACCAAGTGGGTGGCGCCAGGGAGGACGTGGAGCTGGCGCTGTGAGCTTCTACGGGCG 2676
QY 61 GGAACGCCCGGTGCTACTGGCTCAAGGACAGGAAGGCCATCCCAAGAGAGCCAGAAGTATGAT 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2677 GGAACGCCCGGTGCTACTGGCTCAAGGACAGGAAGGCCATCCCAAGAGAGCCAGAAGTATGAT 2736
QY 121 GTGCTCTGCGAGGGGACAGATGGCCATGTGCTGTCATCTCCGCGGGGCTCTGCTCAAGAGACGCG 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2737 GTGCTCTGCGAGGGGACAGATGGCCATGTGCTGTCATCTCCGCGGGGCTCTGCTCAAGAGACGCG 2796
QY 181 GCGAGTACACGTGTGAGGTGGAGGCTTCCAAAGAGCACAGCCAGGCTCCATGTGGAAGAA 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2797 GCGAGTACACGTGTGAGGTGGAGGCTTCCAAAGAGCACAGCCAGGCTCCATGTGGAAGAA 2856
QY 241 AAAGCAAACTGTTTACAGAGGAGCTGACCAATCTGCAGGTGGAGGAGAAAGGCGACAGCT 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2857 AAAGCAAACTGTTTACAGAGGAGCTGACCAATCTGCAGGTGGAGGAGAAAGGCGACAGCT 2916
QY 301 GTGTTTCACTGCAAGACGGAGCACCCCGGGGCCACAGTGACCTGGCGCAAGGGGCTCTTTG 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2917 GTGTTTCACTGCAAGACGGAGCACCCCGGGGCCACAGTGACCTGGCGCAAGGGGCTCTTTG 2976
QY 361 GAGTACGGGCTCTCAGGGAAGCACAGGCCAGCCAGGAGGG 401
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
2977 GAGTACGGGCTCTCAGGGAAGCACAGGCCAGCCAGGAGGG 3017

RESULT 3

US-10-093-463-73
; Sequence 73, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:

; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glennda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Baha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: No. US20030208039A1 Antibodies that Bind to Antigenic Polypepti
; TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.

FILE REFERENCE: 21402-290A (Cura 590AT)
CURRENT APPLICATION NUMBER: US/10/093,463
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/330,380
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/288,342
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/275,578
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 370
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 73
LENGTH: 14061
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (15) ..(14039)
US-10-093-463-73

Query Match 100.0%; Score 401; DB 17; Length 14061;
Best Local Similarity 100.0%; Pred. No. 1.2e-108;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAAGACCAAGTGGTGGCCGACGAGGAGGAGCGTGGAGCTGGCTGTGAGCTGTACGGGCG 60
Db 732 GAAGACCAAGTGGTGGCCGACGAGGAGGAGCGTGGAGCTGGCTGTGAGCTGTACGGGCG 791
Qy 61 GGAACGCCCGTGCACTGGCTGAAGCAGCAGGAGGCGCATCCGCAAGCAGCAGGAAGTATGAT 120
Db 792 GGAACGCCCGTGCACTGGCTGAAGCAGCAGGAGGCGCATCCGCAAGCAGCAGGAAGTATGAT 851
Qy 121 GTGCTTCGAGGGGACGATGCCCATGTGCTCATCCGCGGGCTCGCTCAAGGACGCG 180
Db 852 GTGGTCTCGAGGGGACGATGCCCATGTGCTCATCCGCGGGCTCGCTCAAGGACGCG 911
Qy 181 GCGGAGTACAGTGTGAGGTGGAGGCTTCCAGAGCAGCAGCGCTCCATGTGGAAGAA 240
Db 912 GCGGAGTACAGTGTGAGGTGGAGGCTTCCAGAGCAGCAGCGCTCCATGTGGAAGAA 971
Qy 241 AAAGCAAACTGCTTCACAGAGAGCTGACCAATCTGCAGGTGGAGGAGAAAGGCACGCT 300
Db 972 AAAGCAAACTGCTTCACAGAGAGCTGACCAATCTGCAGGTGGAGGAGAAAGGCACGCT 1031

Qy 301 GTGTTCACTGCAAGACGAGGACCCCGGGCCACAGTACCTGGCGCAAGGGCTCTTGG 360
Db 1032 GTGTTCACTGCAAGACGAGGACCCCGGGCCACAGTACCTGGCGCAAGGGCTCTTGG 1091
Qy 361 GAGCTACGGGCTCAGGGAAGCACCAGCCAGCCAGCCAGGAGGG 401
Db 1092 GAGCTACGGGCTCAGGGAAGCACCAGCCAGCCAGGAGGG 1132

RESULT 4

US-10-093-463-71
Sequence 71, Application US/10093463
Publication No. US20030208039A1
GENERAL INFORMATION:
APPLICANT: Padigaru, Muralidhara
APPLICANT: Shenoy, Suresh
APPLICANT: Kekuda, Rameesh
APPLICANT: Gusev, Vladimir
APPLICANT: Pochart, Pascal
APPLICANT: Zhong, Mei
APPLICANT: Rastelli, Luca
APPLICANT: Mezes, Peter
APPLICANT: Smithson, Glennda
APPLICANT: Guo, Xiaojia
APPLICANT: Gerlach, Valerie
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
APPLICANT: Li, Li
APPLICANT: Zerhusen, Bryan
APPLICANT: Tchernev, Velizar
APPLICANT: Gangolli, Esna
APPLICANT: Vernet, Corine
APPLICANT: Pena, Carol
APPLICANT: Burgess, Catherine
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly
APPLICANT: Gorman, Linda
APPLICANT: Spaderna, Steven
APPLICANT: Voss, Edward
APPLICANT: Malyankar, Uriel
APPLICANT: Anderson, David
APPLICANT: Patturejan, Meera
APPLICANT: Miller, Charles
APPLICANT: Taupier, Raymond J. Jr.
TITLE OF INVENTION: NO. US20030208039A1el Antibodies that Bind to Antigenic Polypepti
FILE REFERENCE: 21402-290A (Cura 590AT)
CURRENT APPLICATION NUMBER: US/10/093,463
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194


```
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 14109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(14088)
US-10-093-463-71

Query Match      100.0%; Score 401; DB 17; Length 14109;
Best Local Similarity 100.0%; Pred. No. 1.2e-108;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAGACCACTGGTGGCGCCGAGGAGGACGCTGGAGCTGCCTGTGAGCTCTCACGGCG 60
Db 732 GAAGACCACTGGTGGCGCCGAGGAGGACGCTGGAGCTGCCTGTGAGCTCTCACGGCG 791

Qy 61 GGAACGCCCGCTGCACTGGCTGAAGACAGGAGGCGCATCCGCAAGAGCCAGCAAGTATGAT 120
Db 792 GGAACGCCCGCTGCACTGGCTGAAGACAGGAGGCGCATCCGCAAGAGCCAGCAAGTATGAT 851

Qy 121 GTGCTCTCGAGGACGACATGCCATGCTGTATCCCGGGGCTCGCTCAAGAGCGCG 180
Db 852 GTGCTCTCGAGGACGACATGCCATGCTGTATCCCGGGGCTCGCTCAAGAGCGCG 911

Qy 181 GCGAGTACAGTGTGAGGTGGAGCTTCCAAAGAGCAGCAGCCTCCATGTGGAAGAA 240
Db 912 GCGAGTACAGTGTGAGGTGGAGCTTCCAAAGAGCAGCAGCCTCCATGTGGAAGAA 971

Qy 241 AAAGCAAACTGCTTTCACAGAGAGCTGACCAATCTGCAAGTGGAGGAAAGGCACAGCT 300
Db 972 AAAGCAAACTGCTTTCACAGAGAGCTGACCAATCTGCAAGTGGAGGAAAGGCACAGCT 1031

Qy 301 GTGTTTCAGTGCAGAGCGGACGACCCCGGCCACAGTACCTGGCGCAAGGGCTCTTG 360
Db 1032 GTGTTTCAGTGCAGAGCGGACGACCCCGGCCACAGTACCTGGCGCAAGGGCTCTTG 1091

Qy 361 GAGCTACGGGCTCAGGGAAGCAGCAGCCAGCCAGCAGGAGG 401
Db 1092 GAGCTACGGGCTCAGGGAAGCAGCAGCCAGCCAGCAGGAGG 1132

RESULT 5
US-10-723-860-5701
; Sequence 5701, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882,0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393

; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5701
; LENGTH: 20489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (565)..(584)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2938)..(2955)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3329)..(3364)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4556)..(4586)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4721)..(4735)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4997)..(5011)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5108)..(5138)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6237)..(6300)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7520)..(7566)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12014)..(12037)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12123)..(12146)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13258)..(13277)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13918)..(13942)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15219)..(15248)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15262)..(15280)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15309)..(15381)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15662)..(15682)
; OTHER INFORMATION: n is a, c, g, or t
```



```
QY 241 AAAGCAAACTGCTTTCACAGAGGAGCTGACCAATCTCGAGTGGAGGAGAAAGGCACAGCT 300
Db 8811 AAAGCAAACTGCTTTCACAGAGGAGCTGACCAATCTCGAGTGGAGGAGAAAGGCACAGCT 8870
QY 301 GTGTTTCAGTGCAGAACGAGGACACCCCGCGGCGCACAGTGAACCTGGCGCAAGGGCCTCTTG 360
Db 8871 GTGTTTCAGTGCAGAACGAGGACACCCCGCGGCGCACAGTGAACCTGGCGCAAGGGCCTCTTG 8930
QY 361 GAGCTACGGGCTCTCAGGAGACACACCCAGCCAGCCAGGAGGG 401
Db 8931 GAGCTACGGGCTCTCAGGAGACACACCCAGCCAGCCAGGAGGG 8971

RESULT 8
US-09-764-891-6045
; Sequence 6045, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6045
; LENGTH: 20565
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-891-6045

Query Match 15.1%; Score 60.6; DB 10; Length 20565;
Best Local Similarity 54.8%; Pred. No. 8.3e-08;
Matches 120; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 10 TGGTGGCGCCAGGGGAGGACGTGAGCTGCGCTGTGAGCTGTCACGGGCGGGAACGCCC 69
Db 13230 TGTGTGGCGCCCTGGGGAGCCAGTGTGCTGAGCTGTGAATGTCTCCGGCTGGCGGCC 13289
QY 70 GTGCACTGGCTGAGGACAGGAAGCCATCCGAGAGCCAGAGATGATGTGCTGTC 129
Db 13290 GTGGTCTGGAGCCACAAATGGAGGCCCTGTGAGAGGCGGAGGCGCTAGAGCTCCATGCC 13349
QY 130 GAGGCGCAGATGGCCATGCTGTATCCCGGGGCTCTGCTCAAGGACGCGGGCGAGTAC 189
Db 13350 GAGGCGCCCGCCGAGTCTCTGATCCAGGCTGCAGGCCCGAGCCCATGCAGGGCTCTAC 13409
QY 190 ACGTGTGAGGTGGAGGCTTCCAGAGCACACCCAGCCCTC 228
Db 13410 ACGTGCCAGTCTGGAGCAGCCCGCGAGCCCAAGCCCTC 13448

RESULT 9
US-10-091-438-270/c
; Sequence 270, Application US/10091438
; Publication No. US20030077606A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ17C1
; CURRENT APPLICATION NUMBER: US/10/091,438
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6045
; LENGTH: 20565
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-438-270/c
```


1	PRIOR APPLICATION NUMBER: 60/229,509
2	PRIOR FILING DATE: 2000-09-05
3	PRIOR APPLICATION NUMBER: 60/236,367
4	PRIOR FILING DATE: 2000-09-29
5	PRIOR APPLICATION NUMBER: 60/237,039
6	PRIOR FILING DATE: 2000-10-02
7	PRIOR APPLICATION NUMBER: 60/237,038
8	PRIOR FILING DATE: 2000-10-02
9	PRIOR APPLICATION NUMBER: 60/236,370
10	PRIOR FILING DATE: 2000-09-29
11	PRIOR APPLICATION NUMBER: 60/236,802
12	PRIOR FILING DATE: 2000-10-02
13	PRIOR APPLICATION NUMBER: 60/237,037
14	PRIOR FILING DATE: 2000-10-02
15	PRIOR APPLICATION NUMBER: 60/237,040
16	PRIOR FILING DATE: 2000-10-02
17	PRIOR APPLICATION NUMBER: 60/240,960
18	PRIOR FILING DATE: 2000-10-20
19	PRIOR APPLICATION NUMBER: 60/239,935
20	PRIOR FILING DATE: 2000-10-13
21	PRIOR APPLICATION NUMBER: 60/239,937
22	PRIOR FILING DATE: 2000-10-13
23	PRIOR APPLICATION NUMBER: 60/241,787
24	PRIOR FILING DATE: 2000-10-20
25	PRIOR APPLICATION NUMBER: 60/246,474
26	PRIOR FILING DATE: 2000-11-08
27	PRIOR APPLICATION NUMBER: 60/246,532
28	PRIOR FILING DATE: 2000-11-08
29	PRIOR APPLICATION NUMBER: 60/249,216
30	PRIOR FILING DATE: 2000-11-17
31	PRIOR APPLICATION NUMBER: 60/249,210
32	PRIOR FILING DATE: 2000-11-17
33	PRIOR APPLICATION NUMBER: 60/226,681
34	PRIOR FILING DATE: 2000-08-22
35	PRIOR APPLICATION NUMBER: 60/225,759
36	PRIOR FILING DATE: 2000-08-14
37	PRIOR APPLICATION NUMBER: 60/225,213
38	PRIOR FILING DATE: 2000-08-14
39	PRIOR APPLICATION NUMBER: 60/227,182
40	PRIOR FILING DATE: 2000-08-22
41	PRIOR APPLICATION NUMBER: 60/225,214
42	PRIOR FILING DATE: 2000-08-14
43	PRIOR APPLICATION NUMBER: 60/235,836
44	PRIOR FILING DATE: 2000-09-27
45	PRIOR APPLICATION NUMBER: 60/230,438
46	PRIOR FILING DATE: 2000-09-06
47	PRIOR APPLICATION NUMBER: 60/215,135
48	PRIOR FILING DATE: 2000-06-30
49	PRIOR APPLICATION NUMBER: 60/225,266
50	PRIOR FILING DATE: 2000-08-14
51	PRIOR APPLICATION NUMBER: 60/249,218
52	PRIOR FILING DATE: 2000-11-17
53	PRIOR APPLICATION NUMBER: 60/249,208
54	PRIOR FILING DATE: 2000-11-17
55	PRIOR APPLICATION NUMBER: 60/249,213
56	PRIOR FILING DATE: 2000-11-17
57	PRIOR APPLICATION NUMBER: 60/249,212
58	PRIOR FILING DATE: 2000-11-17
59	PRIOR APPLICATION NUMBER: 60/249,244
60	PRIOR FILING DATE: 2000-11-17
61	PRIOR APPLICATION NUMBER: 60/249,217
62	PRIOR FILING DATE: 2000-11-17
63	PRIOR APPLICATION NUMBER: 60/249,211
64	PRIOR FILING DATE: 2000-11-17
65	PRIOR APPLICATION NUMBER: 60/249,215
66	PRIOR FILING DATE: 2000-11-17
67	PRIOR APPLICATION NUMBER: 60/249,264
68	PRIOR FILING DATE: 2000-11-17
69	PRIOR APPLICATION NUMBER: 60/249,214

;	PRIOR FILING DATE:	2000-11-17	
;	PRIOR APPLICATION NUMBER:	60/249,297	
;	PRIOR FILING DATE:	2000-11-17	
;	PRIOR APPLICATION NUMBER:	60/232,400	
;	PRIOR FILING DATE:	2000-09-14	
;	PRIOR APPLICATION NUMBER:	60/231,242	
;	PRIOR FILING DATE:	2000-09-08	
;	PRIOR APPLICATION NUMBER:	60/232,081	
;	PRIOR FILING DATE:	2000-09-08	
;	PRIOR APPLICATION NUMBER:	60/232,080	
;	PRIOR FILING DATE:	2000-09-08	
;	PRIOR APPLICATION NUMBER:	60/231,414	
;	PRIOR FILING DATE:	2000-09-08	
;	PRIOR APPLICATION NUMBER:	60/231,244	
;	PRIOR FILING DATE:	2000-09-08	
;	PRIOR APPLICATION NUMBER:	60/233,064	
;	PRIOR FILING DATE:	2000-09-14	
;	PRIOR APPLICATION NUMBER:	60/233,063	
;	PRIOR FILING DATE:	2000-09-14	
;	PRIOR APPLICATION NUMBER:	60/232,397	
;	PRIOR FILING DATE:	2000-09-14	
;	PRIOR APPLICATION NUMBER:	60/232,399	
;	PRIOR FILING DATE:	2000-09-14	
;	PRIOR APPLICATION NUMBER:	60/232,401	
;	PRIOR FILING DATE:	2000-09-14	
;	PRIOR APPLICATION NUMBER:	60/241,808	
;	PRIOR FILING DATE:	2000-10-20	
;	PRIOR APPLICATION NUMBER:	60/241,826	
;	PRIOR FILING DATE:	2000-10-20	
;	PRIOR APPLICATION NUMBER:	60/241,786	
;	PRIOR FILING DATE:	2000-10-20	
;	PRIOR APPLICATION NUMBER:	60/241,221	
;	PRIOR FILING DATE:	2000-10-20	
;	PRIOR APPLICATION NUMBER:	60/246,475	
;	PRIOR FILING DATE:	2000-11-08	
;	PRIOR APPLICATION NUMBER:	60/231,243	
;	PRIOR FILING DATE:	2000-09-08	

Query Match 15.1%; Score 60.6; DB 14; Length 20565;

Query Match 13.1%; Score 00.0; SD 1.1
Best Local Similarity 54.8%; Pred. No. 8.3e-08;

Best local similarity 34.8%; Pseud. NO. 0.35-00;
Matches 120; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

10 TGGGTGGCGCCAGGGAGGACGTGGAGCTGCGCTGTGAGCTGTCA CGGGCGGGAACGCC 69

7336 TGTCTGCGCCCTGCGGAGCCAGCTGCTGCTGAGCTGTGA ACTGTCCCGGGTGGCGCCCG 7277

0
1
2
3
4
5
6
7
8
9
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

[illegible][illegible]

Age Group	Number of people
130-140	1
140-150	2
150-160	3
160-170	4
170-180	5
180-190	6
190-200	7

DB /Z16 GAGGGCCCCCGAGTCCCTCTGCAIACCAGGCCTGCAGGCCCAAG

QY 190 ACGTGTGAGGTGGAGGCTTCCAGAGACACAGCCAGCCCTC 228

Dd

DECEMBER 10

RESULT 10
IIS-09-759-509B-1

US-09-759-508B-1
: Sequence 1 Applicant; Sequence 1, Application US/0
: Publication No. US2002018259

; Publication No. US20020182599A1
; GENERAL INFORMATION:

; GENERAL INFORMATION:
; APPLICANT: Fishman, Mark C.

APPLICANT: FISHMAN, MARK C.
TITLE OF INVENTION: Methods for Diagnosing

; TITLE OF INVENTION: METHODS FOR DIAGNOSIS
; FILE REFERENCE: 00786/381002

FILE REFERENCE: 00786/381002
CURRENT APPLICATION NUMBER: US/09/759,5

; CURRENT AFFILIATION NUMBER: 03/03/155,5
 ; CURRENT FILING DATE: 2001-01-12

; PRIOR APPLICATION NUMBER: US 60/175,787

; PRIOR AFFIDAVIT NUMBER: 00-0
 ;
 ; PRIOR FILING DATE: 2000-01-12


```
; SEQ ID NO 1
; LENGTH: 81940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)..(80910)
; OTHER INFORMATION:
US-09-759-508B-1

Query Match
  14.9%; Score 59.8; DB 9; Length 81940;
Best Local Similarity 49.5%; Pred. No. 1.7e-07;
Matches 154; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 1 GAAGACCACTGGTGGCCAGCGGAGGAGCGTGGAGCTCGCTGTGAGCTGTCCAGGGCG 60
Db 17515 GAGGACCACTGGTGGCCAGCGGAGGAGCGTGGAGCTCGCTGTGAGCTGTCCAGGAA 17574
QY 61 GGAACGCCCGTGCACCTGGCTGAAGGACAGGAAGGCCATCCGCAAGAGCCAGAAAGTATGAT 120
Db 17575 AATGCTAAGGTGAATGGTTCAAAAATGGGACAGAAATCCTCAAAAGCAAGAAATGATGAA 17634
QY 121 GTGGTCTGCGAGGACGATGCCATCTGTGTCATCCGGGGCGCTCGCTCAAGGACCGG 180
Db 17635 AATGTTGCTGATGGCAGGCTCAGAAACTTGTATACATGACTGTACCCAGAGGATATT 17694
QY 181 GCGGAGTACACGTGTGAGGTGGAGGCTTCCAGAGCACAGCCAGCCTCCATGTGGAAGAA 240
Db 17695 AAAACATACACTTGTGATGCTAAGGATTTTAAGACTTCTGTAACTGAATGCTGCTGCCT 17754
QY 241 AAAGCAAACTCTTCCACAGAGGAGCTGACCAATCTGAGGTGGAGGAGAAAGGCACAGCT 300
Db 17755 CCTCATGTGGAAATCTTTAAGACCACTCACCGACCTTCAAGTTAGAGAAAAAGAAATGGCT 17814
QY 301 GTGTTCAAGTG 311
Db 17815 CGATTTGAGTG 17825

RESULT 11
US-09-960-706-1092
; Sequence 1092, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-01US
; CURRENT APPLICATION NUMBER: US/09/960,706
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 60/223,323
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: 09/873,319
; PRIOR FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 1124
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1092
; LENGTH: 81940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134280A1 X90568
US-09-960-706-1092

Query Match
  14.9%; Score 59.8; DB 10; Length 81940;
Best Local Similarity 49.5%; Pred. No. 1.7e-07;
Matches 154; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 1 GAAGACCACTGGTGGCCAGCGGAGGAGCGTGGAGCTCGCTGTGAGCTGTCCAGGGCG 60
Db 17515 GAGGACCACTGGTGGCCAGCGGAGGAGCGTGGAGCTCGCTGTGAGCTGTCCAGGAA 17574
QY 61 GGAACGCCCGTGCACCTGGCTGAAGGACAGGAAGGCCATCCGCAAGAGCCAGAAAGTATGAT 120
Db 17575 AATGCTAAGGTGAATGGTTCAAAAATGGGACAGAAATCCTCAAAAGCAAGAAATGATGAA 17634
QY 121 GTGGTCTGCGAGGACGATGCCATCTGTGTCATCCGGGGCGCTCGCTCAAGGACCGG 180
Db 17635 AATGTTGCTGATGGCAGGCTCAGAAACTTGTATACATGACTGTACCCAGAGGATATT 17694
QY 181 GCGGAGTACACGTGTGAGGTGGAGGCTTCCAGAGCACAGCCAGCCTCCATGTGGAAGAA 240
Db 17695 AAAACATACACTTGTGATGCTAAGGATTTTAAGACTTCTGTAACTGAATGCTGCTGCCT 17754
QY 241 AAAGCAAACTCTTCCACAGAGGAGCTGACCAATCTGAGGTGGAGGAGAAAGGCACAGCT 300
Db 17755 CCTCATGTGGAAATCTTTAAGACCACTCACCGACCTTCAAGTTAGAGAAAAAGAAATGGCT 17814
QY 301 GTGTTCAAGTG 311
Db 17815 CGATTTGAGTG 17825

RESULT 12
US-09-873-319-724
; Sequence 724, Application US/09873319A
; Publication No. US20030134324A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; APPLICANT: Kulkarni, Prakash.
; APPLICANT: Getzenberg, Robert H.
; APPLICANT: Waga, Iwao
; APPLICANT: Yamamoto, Jun
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; FILE REFERENCE: 44921-5029-US
; CURRENT APPLICATION NUMBER: US/09/873,319A
; CURRENT FILING DATE: 2001-06-05
; EARLIER APPLICATION NUMBER: US 60/223,323
; EARLIER FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 755
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 724
; LENGTH: 81940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20030134324A1 X90568
US-09-873-319-724

Query Match
  14.9%; Score 59.8; DB 10; Length 81940;
Best Local Similarity 49.5%; Pred. No. 1.7e-07;
Matches 154; Conservative 0; Mismatches 157; Indels 0; Gaps 0;

QY 1 GAAGACCACTGGTGGCCAGCGGAGGAGCGTGGAGCTCGCTGTGAGCTGTCCAGGGCG 60
Db 17515 GAGGACCACTGGTGGCCAGCGGAGGAGCGTGGAGCTCGCTGTGAGCTGTCCAGGAA 17574
QY 61 GGAACGCCCGTGCACCTGGCTGAAGGACAGGAAGGCCATCCGCAAGAGCCAGAAAGTATGAT 120
Db 17575 AATGCTAAGGTGAATGGTTCAAAAATGGGACAGAAATCCTCAAAAGCAAGAAATGATGAA 17634
QY 121 GTGGTCTGCGAGGACGATGCCATCTGTGTCATCCGGGGCGCTCGCTCAAGGACCGG 180
Db 17635 AATGTTGCTGATGGCAGGCTCAGAAACTTGTATACATGACTGTACCCAGAGGATATT 17694
QY 181 GCGGAGTACACGTGTGAGGTGGAGGCTTCCAGAGCACAGCCAGCCTCCATGTGGAAGAA 240
Db 17695 AAAACATACACTTGTGATGCTAAGGATTTTAAGACTTCTGTAACTGAATGCTGCTGCCT 17754
QY 241 AAAGCAAACTCTTCCACAGAGGAGCTGACCAATCTGAGGTGGAGGAGAAAGGCACAGCT 300
Db 17755 CCTCATGTGGAAATCTTTAAGACCACTCACCGACCTTCAAGTTAGAGAAAAAGAAATGGCT 17814
QY 301 GTGTTCAAGTG 311
Db 17815 CGATTTGAGTG 17825
```


Qy 241 AAAGCRAACTGCTTTCACAGAGGAGCTGACCAATCTGCAGGTGGAGGAGAAAGGCACAGCT 300
Db 17846 CCTCATGTGGAAATCTTAAGACCACCTCACCGACCTTCAGTTAGAGAAAAGAAATGGCT 17905
Qy 301 GTGTTACAGTG 311
Db 17906 CGATTTGAGTG 17916

Search completed: March 21, 2005, 16:07:56
Job time : 258.197 secs

GenCore version 5.1.6
Copyright (C) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:32:32 ; Search time 689.694 Seconds
(without alignments)
9503.485 Million cell updates/sec

Title: US-10-077-130-6_COPY_13600_14700

Perfect score: 1101
Sequence: 1 cacacagccacactgtgac.....agcaggaaggcccatgttc 1101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1101	100.0	23907	13	US-10-077-130-6
2	1101	100.0	24120	13	US-10-077-130-4
3	1076	97.7	20489	18	US-10-723-860-5701
4	886.8	80.5	14061	17	US-10-093-483-73
5	886.8	80.5	14109	17	US-10-093-483-71
6	53.4	4.9	2534	17	US-10-108-260A-1946
7	53.4	4.9	7564	18	US-10-476-397-15
8	48.6	4.4	1813	18	US-10-723-860-5479
9	46.6	4.2	3613	18	US-10-723-860-4453
10	46.6	4.2	4176	18	US-10-723-860-8268
11	45.6	4.1	1529	18	US-10-437-963-48468

12	45	4.1	27705	10	US-09-942-025-12	Sequence 12, Appl
13	45	4.1	67311	10	US-09-942-025-1	Sequence 1, Appl
14	44.4	4.0	1469	18	US-10-437-963-64043	Sequence 1, Appl
15	44.2	4.0	4698	15	US-10-156-761-6923	Sequence 6923, Ap
16	44.2	4.0	9025608	15	US-10-156-761-1	Sequence 1, Appl
17	43.8	4.0	815	16	US-10-257-836A-255	Sequence 255, App
18	43.8	4.0	817	16	US-10-257-836A-252	Sequence 252, App
19	43.8	4.0	819	16	US-10-257-836A-251	Sequence 251, App
20	43.8	4.0	820	16	US-10-257-836A-253	Sequence 253, App
21	43.8	4.0	825	16	US-10-257-836A-254	Sequence 254, App
22	43.8	4.0	1038	17	US-10-369-493-31577	Sequence 31577, A
23	43.8	4.0	1725	15	US-10-156-761-2800	Sequence 2800, Ap
24	43.8	4.0	9590	17	US-10-282-132A-33351	Sequence 33351, A
25	43.8	4.0	177587	13	US-10-087-192-1438	Sequence 1438, Ap
26	43.8	4.0	9025608	15	US-10-156-761-1	Sequence 1, Appl
27	43.4	3.9	1338	17	US-10-685	Sequence 35, Appl
28	43.4	3.9	1599	18	US-10-437-963-36356	Sequence 36356, A
29	43.2	3.9	1578	16	US-10-029-386-20397	Sequence 20397, A
30	43.2	3.9	3340	16	US-10-251-186-20	Sequence 20, Appl
31	43.2	3.9	3340	17	US-10-291-172-168	Sequence 168, App
32	43.2	3.9	3340	17	US-10-221-278-168	Sequence 168, App
33	43	3.9	6561	18	US-10-437-963-9359	Sequence 9359, Ap
34	42.8	3.9	2454	15	US-10-156-761-6929	Sequence 6929, Ap
35	42.6	3.9	1661	18	US-10-723-860-8081	Sequence 8081, Ap
36	42.6	3.9	2864	18	US-10-437-963-101596	Sequence 101596, A
37	42.4	3.9	474	18	US-10-425-115-88598	Sequence 88598, A
38	42.4	3.9	1192	17	US-10-260-238-3989	Sequence 3989, Ap
39	42.4	3.9	1580	18	US-10-437-963-88464	Sequence 88464, A
40	42.4	3.9	1723	17	US-10-425-114-24896	Sequence 24896, A
41	42.4	3.9	1737	17	US-10-425-114-16977	Sequence 16977, A
42	42.2	3.8	648	18	US-10-437-963-17286	Sequence 17286, A
43	42.2	3.8	1789	18	US-10-437-963-101196	Sequence 101196, A
44	42	3.8	659	16	US-10-058-597-29	Sequence 29, Appl
45	42	3.8	980	18	US-10-437-963-10642	Sequence 10642, A

ALIGNMENTS

RESULT 1

US-10-077-130-6
; Sequence 6, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047PINCPI(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 23907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-077-130-6

Query Match 100.0%; Score 1101; DB 13; Length 23907;
Best Local Similarity 100.0%; Pred. No. 1.9e-276;
Matches 1101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CACACAGCCACACTGTGACACTGTCTGGGAGCTCCCATGATGATGAGCGGTGT 60
Db 13600 CACACAGCCACACTGTGACACTGTCTGGGAGCTCCCATGATGATGAGCGGTGT 13659
Qy 61 CTCTGTGGTACCGGTGGAGGTGAAGAGGGGGCCACAGGCCAGTGGCGGTGTGCCAC 120
Db 13660 CTCTGTGGTACCGGTGGAGGTGAAGAGGGGGCCACAGGCCAGTGGCGGTGTGCCAC 13719

Db 13797 GAGCTGTGCTGTGACCCGAGTGTGTGTGATGCTGGCCCTGGCCCGGGGAGACCTACCCG 13856
QY 181 TTCCGTGTGGCAGCTGTGGGCTCTGTGGGTGCTGGGGAACCGGTTACCTGCCCGCCAGACA 240
Db 13857 TTCCGTGTGGCAGCTGTGGGCTCTGTGGGTGCTGGGGAACCGGTTACCTGCCCGCCAGACA 13916
QY 241 GTGCGGCTTGAGAGCCACCGAAGCCCTGTGCTCTCCAGCCCTCAGCCCTCAGAGCCCG 300
Db 13917 GNN 13976
QY 301 CAGGTGGCAGCTGGTGAAGATGTCTCTGTGAGCTGTGAGTGTGCTGAGGCTGTGTGAG 360
Db 13977 CAGGTGGCAGCTGGTGAAGATGTCTCTGTGAGCTGTGAGTGTGCTGAGGCTGTGTGAG 14036
QY 361 GTCATCTGGCACAGGAATGAGCGCATCCAGCCGCTGGGCGGTTCCGAGTGTGCTCC 420
Db 14037 GTCATCTGGCACAGGAATGAGCGCATCCAGCCGCTGGGCGGTTCCGAGTGTGCTCC 14096
QY 421 CAGGTGGCAACAGATGTGTGATCAAGGGCTTCAAGGCAAGAACACGAGGCGAGTAC 480
Db 14097 CAGGTGGCAACAGATGTGTGATCAAGGGCTTCAAGGCAAGAACACGAGGCGAGTAC 14156
QY 481 CACTGTGGCTGTGCTCAGGGCTCCATCTGCCCTGGCGCTGCCACTTCCAGTGTGCACTG 540
Db 14157 CACTGTGGCTGTGCTCAGGGCTCCATCTGCCCTGGCGCTGCCACTTCCAGTGTGCACTG 14216
QY 541 AGCCCGACCTCTGTGATGAGGCCCTCAGCCAGCTTGGCCCCCGGAGCAGCCAGGAG 600
Db 14217 AGCCCGACCTCTGTGATGAGGCCCTCAGCCAGCTTGGCCCCCGGAGCAGCCAGGAG 14276
QY 601 GGTGACCTGCACCTACTGTGGAGGCCCTGGCTCGGAAACGTGCGATGAGCCGTGAGCCC 660
Db 14277 GGTGACCTGCACCTACTGTGGAGGCCCTGGCTCGGAAACGTGCGATGAGCCGTGAGCCC 14336
QY 661 ACCTGCTGCTCCATTAGCAGCTGCCAGAGGAGCGCGCTCGCAGCGCTGCCAG 720
Db 14337 ACCTGCTGCTCCATTAGCAGCTGCCAGAGGAGCGCGCTCGCAGCGCTGCCAG 14396
QY 721 GAGGCAGAGAGGTGCACTGTATCTCTGAGGCTACTCCACGCGCGATGAGCTGGCC 780
Db 14397 GAGGCAGAGAGGTGCACTGTATCTCTGAGGCTACTCCACGCGCGATGAGCTGGCC 14456
QY 781 CCGACTGGAGATGCTGACCTCTCACACACAGCTGTGATGATGATGCCGGGCGAGCCAC 840
Db 14457 CCGACTGGAGATGCTGACCTCTCACACACAGCTGTGATGATGATGCCGGGCGAGCCAC 14516
QY 841 CTTTCCTGCTCACTTACCTCAAGAGGCTGGGAGCGCAGGCACTTCACTGCTGCGCCAGC 900
Db 14517 CTTTCCTGCTCACTTACCTCAAGAGGCTGGGAGCGCAGGCACTTCACTGCTGCGCCAGC 14576
QY 901 AAGTTTGGGGCCCGAGAGCCCTCTGTGAAGCCACAGCAGCAGCAGGAGCCACTGGCT 960
Db 14577 AAGTTTGGGGCCCGAGAGCCCTCTGTGAAGCCACAGCAGCAGCAGGAGCCACTGGCT 14636
QY 961 GCTGTGCGCCACCACTGGGAGCTGAGCACAAGACCTGGGTGATCCCTCAATGAC 1020
Db 14637 GCTGTGCGCCACCACTGGGAGCTGAGCACAAGACCTGGGTGATCCCTCAATGAC 14696
QY 1021 AAGCAGCTGTGAAGATCAGGCTGCTTTTAAGGCTTCAAGGTCCGGAAGAGATGAAG 1080
Db 14697 AAGCAGCTGTGAAGATCAGGCTGCTTTTAAGGCTTCAAGGTCCGGAAGAGATGAAG 14756
QY 1081 CAGCAGGAAGGCCCATGTTTC 1101
Db 14757 CAGCAGGAAGGCCCATGTTTC 14777

RESULT 4

US-10-093-463-73

; Sequence 73, Application US/10093463

; Publication No. US20030208039A1

; GENERAL INFORMATION:

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glennnda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Tchernus, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: No. US20030208039A1el Antibodies that Bind to Antigenic Polypeptide
; FILE OF INVENTION: Encoding The Antigens, and Methods of Use.
; FILE REFERENCE: 21402-290A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73

; LENGTH: 14061
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(14039)
US-10-093-463-73

Query Match 80.5%; Score 886.8; DB 17; Length 14061;
Best Local Similarity 98.9%; Pred. No. 9.2e-221;
Matches 904; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 1 CACAGCAGCACACTGTGACACTGTCTTGGGAGCTCCCATGATGTGGAGCGGTGT 60
Db |||||
QY 9366 CGCAGCAGCACACTGTGACACTGTCTTGGGAGCTCCCATGATGTGGAGCGGTGT 9425
Db |||||
QY 61 CTCTGTGGCTACCGGTGTGAGGTGAAGAGGGGGCCACAGGCCAGTGGCGGTGTGCCAC 120
Db |||||
QY 9426 CTCTGTGGCTACCGGTGTGAGGTGAAGAGGGGGCCACAGGCCAGTGGCGGTGTGCCAC 9485
Db |||||
QY 121 GAGCTGTGTGACCGGAGTGTGTGTGGATGGCTGGCCCCGGGGAGACCTACCGC 180
Db |||||
QY 9486 GAGCTGTGTGACCGGAGTGTGTGTGGATGGCTGGCCCCGGGGAGACCTACCGC 9545
Db |||||
QY 181 TTCCGTGTGGCAGCTGTGGGCCCTGTGGTGTGGGAAACCGGTTCACTGCCCCAGACA 240
Db |||||
QY 9546 TTCCGTGTGGCAGCTGTGGGCCCTGTGGTGTGGGAAACCGGTTCACTGCCCCAGACA 9605
Db |||||
QY 241 GTCCGGCTTGACAGAGCCACCGAAGCTGTGCTCTCCCGAGCCCTCAGCCCTCAGAGCCGG 300
Db |||||
QY 9606 GTCCGGCTTGACAGAGCCACCGAAGCTGTGCTCTCCCGAGCCCTCAGCCCTCAGAGCCGG 9662
Db |||||
QY 301 CAGGTGGCAGCTGTGTAAGATGTCTCTGGAGCTTGAGGTGGTGGCTGAGGCTGGTGAG 360
Db |||||
QY 9663 CAGGTGGCAGCTGTGTAAGATGTCTCTGGAGCTTGAGGTGGTGGCTGAGGCTGGTGAG 9722
Db |||||
QY 361 GTCACTGTGCAACAGGGAATGAGCGCATCCAGCCCGGTGGCGGTTGAGGTGTCTCC 420
Db |||||
QY 9723 GTCACTGTGCAACAGGGAATGAGCGCATCCAGCCCGGTGGCGGTTGAGGTGTCTCC 9782
Db |||||
QY 421 CAGGCTGGCAACAGATGCTGTGTATCAAGGCTTTCAGGGCAGAGACCCAGGCGGAGTAC 480
Db |||||
QY 9783 CAGGCTGGCAACAGATGCTGTGTATCAAGGCTTTCAGGGCAGAGACCCAGGCGGAGTAC 9842
Db |||||
QY 481 CACTGTGGCTTGGCTCAGGGTCCATCTGCTCCCTGGGCTGGCCACTTTCAGGTGGCACTG 540
Db |||||
QY 9843 CACTGTGGCTTGGCTCAGGGTCCATCTGCTCCCTGGGCTGGCCACTTTCAGGTGGCACTG 9902
Db |||||
QY 541 AGCCGAGCTCTGTGGATGAGCGCCCTCAGCCAGCTTGGCCCCCGAGGCGAGCCAGGAG 500
Db |||||
QY 9903 AGCCGAGCTCTGTGGATGAGCGCCCTCAGCCAGCTTGGCCCCCGAGGCGAGCCAGGAG 9962
Db |||||
QY 601 GGTGACCTGACCTACTGTGGAGGCTTGGCTCGGAAACGTCGATGAGCGGTGAGGCC 660
Db |||||
QY 9963 GGTGACCTGACCTACTGTGGAGGCTTGGCTCGGAAACGTCGATGAGCGGTGAGGCC 10022
Db |||||
QY 661 AGCTGGACTCCATTAGGAGCTGCCAGAGGAGCGGCCGCTCGAGCGCTTGGCCACAG 720
Db |||||
QY 10023 AGCTGGACTCCATTAGGAGCTGCCAGAGGAGCGGCCGCTCGAGCGCTTGGCCACAG 10082
Db |||||
QY 721 GAGGAGAGGAGGTGGCACTGATCTCTCTGAAGGCTACTCAGCGCCGATGAGCTGGCC 780
Db |||||
QY 10083 GAGGAGAGGAGGTGGCACTGATCTCTCTGAAGGCTACTCAGCGCCGATGAGCTGGCC 10142
Db |||||
QY 781 CGCACTGGAGATGCTGACCTCTCACACACAGCTCTGATGATGAGTCCCGGGCAGGCAAC 840
Db |||||
QY 10143 CGCACTGGAGATGCTGACCTCTCACACACAGCTCTGATGATGAGTCCCGGGCAGGCAAC 10202
Db |||||
QY 841 CCTTCCCTGGTCACTTACCTCAAGAGGCTGGAGGCCAGGCACTTCACTTGGCCAGC 900
Db |||||
QY 10203 CCTTCCCTGGTCACTTACCTCAAGAGGCTGGAGGCCAGGCACTTCACTTGGCCAGC 10262
Db |||||
QY 901 AAGTTGGGGCCCC 914

Db 10263 AAGTTGGGGCCCC 10276

RESULT 5
US-10-093-463-71
; Sequence 71, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Fascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glennnda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Pattursajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: NO. US20030208039A1el Antibodies that Bind to Antigenic Polypepti
; FILE REFERENCE: 21402-290A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03
; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380


```
/ PRIOR FILING DATE: 2001-10-18
/ PRIOR APPLICATION NUMBER: 60/275,235
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: 60/288,342
/ PRIOR FILING DATE: 2001-05-03
/ PRIOR APPLICATION NUMBER: 60/275,578
/ PRIOR FILING DATE: 2001-03-13
/ NUMBER OF SEQ ID NOS: 370
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 71
/ LENGTH: 14109
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (15)..(14088)
US-10-093-463-71

Query Match      80.5%; Score 886.8; DB 17; Length 14109;
Best Local Similarity 98.9%; Pred. No. 9.2e-221;
Matches 904; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY      1  CACAGCAGCCACTGTGACACTGTCTTGGCAGCTCCCATGAGTGTGAGGCGGTGT 60
Db      9366 CGCAGCAGCCACTGTGACACTGTCTTGGCAGCTCCCATGAGTGTGAGGCGGTGT 9425

QY      61  CTCTGTGGCTACCGCTGGAGGTGAAGAGGGGGCCACAGGCCAGTGGGGCTGTGCCAC 120
Db      9426 CTCTGTGGCTACCGCTGGAGGTGAAGAGGGGGCCACAGGCCAGTGGGGCTGTGCCAC 9485

QY      121 GAGCTGGTGGACCCGAGTGTGTGTGATGTGCTGGCCCCCGGGAGACCTTACC 180
Db      9486 GAGCTGGTGGACCCGAGTGTGTGTGATGTGCTGGCCCCCGGGAGACCTTACC 9545

QY      181 TTCCTGTGCGAGCTGTGGCCCTGTGGTGTCTGGGGAAACCGTTTCACTTGC 240
Db      9546 TTCCTGTGCGAGCTGTGGCCCTGTGGTGTCTGGGGAAACCGTTTCACTTGC 9605

QY      241 GTGCGGCTTTCAGAGCACCAGAGCCTGTGCTCTCCAGCCCTCAGCCCTGAGAGCCGG 300
Db      9606 GTGCGGCTTTCAGAGCACCAGAGCCTGTGCTCTCCAGCCCTCAGCCCTGAGAGCCGG 9662

QY      301 CAGGTGGCAGCTGTGTAAGATGTCTCTGAGAGCTGAGTGGTGGCTGAGGCTGGT 360
Db      9663 CAGGTGGCAGCTGTGTAAGATGTCTCTGAGAGCTGAGTGGTGGCTGAGGCTGGT 9722

QY      361 GTCTCTGSCACAAGGGAATGAGCGCATCCAGCCCGGTGGCGGTTTGAGGTGTCTCC 420
Db      9723 GTCTCTGSCACAAGGGAATGAGCGCATCCAGCCCGGTGGCGGTTTGAGGTGTCTCC 9782

QY      421 CAGGGTTCGSCAACAGATGCTGGTGTATCAAGGGCTTTCAGCGCAGAGACACAGGGCGAGTAC 480
Db      9783 CAGGGTTCGSCAACAGATGCTGGTGTATCAAGGGCTTTCAGCGCAGAGACACAGGGCGAGTAC 9842

QY      481 CACTGTGGCTGGCTCAGGGCTCCATCTGCCCCCTGGCTGGCCACCTTCCAGGTGGCACTG 540
Db      9843 CACTGTGGCTGGCTCAGGGCTCCATCTGCCCCCTGGCTGGCCACCTTCCAGGTGGCACTG 9902

QY      541 AGCCCGAGCCTCTGTGGATGAGGCCCTCAGCCCGCTTGGCCCCCGGAGCCAGCCAGGAG 600
Db      9903 AGCCCGAGCCTCTGTGGATGAGGCCCTCAGCCCGCTTGGCCCCCGGAGCCAGCCAGGAG 9962

QY      601 GGTGACCTGACCTACTGTGGGAGCCCTGTGCTCGAAACGTTCGATGAGCCGTGAGCCC 660
Db      9963 GGTGACCTGACCTACTGTGGGAGCCCTGTGCTCGAAACGTTCGATGAGCCGTGAGCCC 10022

QY      661 AGCTTGGACTCCATTAGGAGCTGCCAGGAGGAGGCGCGCTCGACGCGCTGCCACAG 720
Db      10023 AGCTTGGACTCCATTAGGAGCTGCCAGGAGGAGGCGCGCTCGACGCGCTGCCACAG 10082

QY      721 GAGGCAGAGGAGGTGGCACCCTGATCTCTCTGAAGGCTACTCTCAGCGCGATGAGCTGGCC 780
Db      10083 GAGGCAGAGGAGGTGGCACCCTGATCTCTCTGAAGGCTACTCTCAGCGCGATGAGCTGGCC 10142
```

```
QY      781  CGCACTGGAGATGCTGACCTCTCACACACACAGCTCTGTATGATGAGTCCCGGGCAGGCACC 840
Db      10143 CGCACTGGAGATGCTGACCTCTCACACACACAGCTCTGTATGATGAGTCCCGGGCAGGCACC 10202

QY      841  CCTTCCCTGGTCACTTACCTCAAGAGGCTGGGAGGCCAGGCACCTCACCACTGGCCAGC 900
Db      10203 CCTTCCCTGGTCACTTACCTCAAGAGGCTGGGAGGCCAGGCACCTCACCACTGGCCAGC 10262

QY      901  AAGTTTGGGCCCC 914
Db      10263 AAGTTGAGCCCCCCC 10276

RESULT 6
US-10-108-260A-1946
; Sequence 1946, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1el full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1946
; LENGTH: 2534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1946

Query Match      4.9%; Score 53.4; DB 17; Length 2534;
Best Local Similarity 53.6%; Pred. No. 0.0004;
Matches 111; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY      287  CCCCTGAGAGCGCGCAGGTGGCAGCTGGTGAAGATGTCTCTTGGAGCTTGAGGTGGTGG 346
Db      803  CCCCTGAAGACCACTGGTGGTGGCGCCAGGGAGGACGTGGAGCTGGCTGTGAGCTGTTCAC 862

QY      347  CTGAGGCTGGTGGAGTCACTCTGGCAAAAGGAATGGAGCGCATCCAGCCCGTGGCGGT 406
Db      863  GGGCGGGAACCGCCGCTGCACTGGCTGAAGGACAGGAAGGCCATCCGCAAGAGCCAGAAAT 922

QY      407  TCGAGGTGTCTCCACGGTGGCGCACAGATGCTGGTGATCAAGGGCTTTCACGCGCACAAG 466
Db      923  ATGATGTGGTCTGGAGGGCAGATGGCCATGCTGGTCTATCCGCGGGGCTTCGCTCAAGG 982

QY      467  ACCAGGCGCAGTACCACTGTGGCTGG 493
Db      983  ACGCGGCGAGTACACGTGTGAGTGG 1009

RESULT 7
US-10-476-397-15
; Sequence 15, Application US/10476397
; Publication No. US20040115687A1
; GENERAL INFORMATION:
; APPLICANT: YUE, Henry
; APPLICANT: LEE, Ernestine A.
; APPLICANT: DUGGAN, Brendan M.
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: HONCHELL, Cynthia D.
; APPLICANT: DING, Li
; APPLICANT: JACKSON, Jennifer L.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: KALLICK, Deborah A.
; APPLICANT: LEE, Sally
; APPLICANT: WARREN, Bridget A.
; APPLICANT: XU, Yuming
; APPLICANT: TRAN, Uyen K.
; APPLICANT: LAL, Preeti G.
; APPLICANT: THORNTON, Michael B.
```



```
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: YAO, Monique G.
; APPLICANT: NGUYEN, Damien B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: KHAN, Farrah A.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: GRIFPIN, Jennifer A.
; APPLICANT: CHINN, Anna M.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: FORSYTHE, Ian J.
; TITLE OF INVENTION: CELL ADHESION AND EXTRACELLULAR MATRIX PROTEINS
; FILE REFERENCE: PF-0968 USN
; CURRENT APPLICATION NUMBER: US/10/476,397
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: PCT/US02/13874
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/288,290
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/292,468
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/298,616
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/301,672
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/345,008
; PRIOR FILING DATE: 2002-01-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 7564
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7326129CB1
US-10-476-397-15

Query Match 4.9%; Score 53.4; DB 18; Length 7564;
Best Local Similarity 53.6%; Pred. No. 0.0004;
Matches 111; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

QY 287 CCCTGAGACCGGAGGTGGGAGCTGGTGAAGATCTCTCTGGAGCTTGAGTGGTGG 346
DB 2612 CCCTGGAAGACCAAGTGGTGGGCGCCAGGAGGACGTGGAGCTGCGCTGTGAC 2671

QY 347 CTGAGCTGGTGAAGTCACTGGCAACAGGGAATGGAGCGCATCCAGCCCGTGGCGGT 406
DB 2672 GGGCGGGAACGCCGTGCACTGGCTGAAGACAGGAGGCCATCCCAAGAGCCAGAGT 2731

QY 407 TCGAGGTGGTCTCCAGGTCGGCAACAGATGCTGGTGATCAAGGGCTTTCACGGCAGAA 466
DB 2732 ATGATGTGGTCTCGGAGGCGACAGTGGCCATGCTGGTCAATCCGCGGGCTCGCTCAAG 2791

QY 467 ACCAGGGCGAGTACCACTGTGGCTGG 493
DB 2792 ACCGGGGCGAGTACACGTGTGAGGTGG 2818

RESULT 8
US-10-723-860-5479
; Sequence 5479, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4453
; LENGTH: 3613
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-4453

Query Match 4.2%; Score 46.6; DB 18; Length 3613;
Best Local Similarity 53.8%; Pred. No. 0.023;
Matches 119; Conservative 0; Mismatches 99; Indels 3; Gaps 1;

QY 6 CAGCCACACTGTGACACTGTCTTGGCAGCTCCCATGATGATGGAGCGGTGGTCTCTG 65
DB 1743 CAGCCAGGGCATCACACTGACATGACACACCTCTCGGGGCCCGGAGCCCATCTCT 1802

QY 66 TGGCTACCGCGTGGAGGTG---AAGAGGGGGGCCACAGCCAGTGGCGGTGTGGCCACGA 122
DB 1803 GGGCTACCTGATCGAGAGCGGTAAAGGGGAGACACCTCGACGGCAGTGNACGCCA 1862
```

```
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 5479
; LENGTH: 1813
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1166)..(1183)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-5479

Query Match 4.4%; Score 48.6; DB 18; Length 1813;
Best Local Similarity 52.2%; Pred. No. 0.007;
Matches 133; Conservative 0; Mismatches 119; Indels 3; Gaps 1;

QY 4 AGCAGCCACACTGTGACACTGTCTTGGGCAGCTCCCATGATGATGGAGCGGTGGTCTC 63
DB 305 AGCAGAGCTCTGTGACTGTGTAGCTGGAGCCCCAGAGAGGCTGGGGAGGCTTGGGCC 364

QY 64 TGTGGCTACCGGTGGAGGT---GAAGAGAGGGGGGCCACAGGCCAGTGGCGGCTGTGCCAC 120
DB 365 CAGGGCTATGTCTGGAGCTCTCGAGAGAGGGAGGCTCTCGAGTGGGTGCTGTGAGTGCC 424

QY 121 GAGCTGGTGGTGGAGCCCGAGTGTGTGTGATGGCTTGGCCCCCGGGGAGACCTTACCGC 180
DB 425 CGGGCCATGATGGTGAACCCAGCAGACTGTGCGGAACCTGGCTCTGGGAGACAAGTTCTC 484

QY 181 TTCCGTGTGGCAGCTGTGGCCCTGTGGTGTCTGGGGAACCGGTTTCACTGCCCCCAGACA 240
DB 485 CTGCGCGTGTCTGCACTGAGTTCTGACGGGGCTTGGCCCCCGGCCCATGCTTGACACGCC 544

QY 241 GTGCGGCTTGCAGAG 255
DB 545 ATCCACATCCGAGAG 559

RESULT 9
US-10-723-860-4453
; Sequence 4453, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: Patentin version 3.2
; SEQ ID NO 4453
; LENGTH: 3613
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-4453

Query Match 4.2%; Score 46.6; DB 18; Length 3613;
Best Local Similarity 53.8%; Pred. No. 0.023;
Matches 119; Conservative 0; Mismatches 99; Indels 3; Gaps 1;

QY 6 CAGCCACACTGTGACACTGTCTTGGCAGCTCCCATGATGATGGAGCGGTGGTCTCTG 65
DB 1743 CAGCCAGGGCATCACACTGACATGACACACCTCTCGGGGCCCGGAGCCCATCTCT 1802

QY 66 TGGCTACCGCGTGGAGGTG---AAGAGGGGGGCCACAGCCAGTGGCGGTGTGGCCACGA 122
DB 1803 GGGCTACCTGATCGAGAGCGGTAAAGGGGAGACACCTCGACGGCAGTGNACGCCA 1862
```



```
; SEQ ID NO 12
; LENGTH: 27705
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-942-025-12

Query Match
Best Local Similarity 4.1%; Score 45; DB 10; Length 27705;
Matches 298; Conservative 0; Mismatches 360; Indels 7; Gaps 3;

QY 74 GCGTGAGGTGAAGGAGGGGCCACAGGCCAGTGGCGGTGTGCCACGAGCTGGTGCCTG 133
DB 26087 GCGCGGTGGCGCGCGCCAGACGACGCGTTCAGGACCTGGCGCGCGCGCTGTGGG 26146
QY 134 GACCCGAGTGTGTGTGATGCTGGCCCGCGGAGACCTACCGCTTCCTCGGTGGCAG 193
DB 26147 GCGTGTTCGCGCGCGCGAAGCGACACCGCGGCTGCGCTTGTGATGATGTGG 26205
QY 194 CTGTGGGCCCTGTGGGTGCTGGGAACCGGTTACCTGCCCGCCAGACAGTGGCGTTCAG 253
DB 26207 GCACCGAGCCGTGACGCTGTGGCTGTGGCGGGCGCTGGCGAGCGCGCGGAGCCGG 26266
QY 254 AGCCACCGAAGCTGTGCTCCCGACCCCTCAGCCCTTGAGAGCGGCGAGGTGGCAGCTG 313
DB 26267 ACCTTGGCGTGGCGGGCGCTGCGCTGGCGCGCGCTGTGGCGCACAGGCGGTAG 26326
QY 314 GTGAAGATGTCTCTGTGAGCTTGTGAGTGTGGTGGCTGAGCTGGTGAAGTTCATCTGGCACA 373
DB 26327 CGGAAGAGCTCACCCGAGCCCGAGCTGGACCTGCGGGCACGGTGTGGTGACCGGG 26386
QY 374 AGGGAATGAGCGATCCAGCCCGGTGGGGTTCGAGGTGTCTCCAGAGGTGCGCAAC 433
DB 26387 GCACAGGGGAGCTGGGTACGGCGGTGCGCGCGACCTGTGTGGCGCGCACGGGGTGGCG 26446
QY 434 AGATGCTGTGTA--TCAGGGGCTTCACGGCAGAGAACAGGCGAGTACCACTGTGGCT 491
DB 26447 ACCTTGTGCTGAGCTGCGCGCGGGCTGAGGCGCCCGGGCGCGAGCTTGTGGAT 26506
QY 492 GCGTCAGGGCTCATCTGCGCTGGCTGCCACTTTCAGGTGGCACTGAGCCCGACCTC 551
DB 26507 CGCTCGCGAGTTCGCGCGCGAGACGGTGAACGTGTGGCGCGCTGCGACGTGTGCAAGC--G 26564
QY 552 TGTGATAGGGCCCTTCAGCCCGAGTTCGCCCGAGGCGAGCCAGGAGGTGACCTGCA 611
DB 26565 GAGAGAGTTCGCGGTGTCTGGCGGATTCAGCGCGCGCGCGCGCTGAGCGCGGTGT 26624
QY 612 CCTACTGTGGAGGCGCTGGCTCGGAACGTCGCGATGAGCCGTGAGCCCGACGCTGACTC 671
DB 26625 GCACCTGGCGCGCGCTGACGACGCGCGTGTCTGCCGCGCAGACGGCGAGCGCTCTC 26684
QY 672 CATTAGCGAGTTCAGAGGAGGACGGCGCTTCGAGCGCTTCGACAGGCGCAGAGGA 731
DB 26685 ---GCGGGTGTGGCGCGCGAAGGTGACGGGGCGCTGCACCTGCACGAGCTGACGCGGA 26741
QY 732 GGTGG 736
DB 26742 GCTGG 26746

RESULT 13
US-09-942-025-1
; Sequence 1, Application US/09942025
; Publication No. US20030054547A1
; GENERAL INFORMATION:
; APPLICANT: Kosan Biosciences, Inc.
; APPLICANT: Julien, Bryan
; TITLE OF INVENTION: POLYKETIDE SYNTHASE GENE FROM SORANGIUM
; FILE REFERENCE: 30062-20020.21
; CURRENT APPLICATION NUMBER: US/09/942,025
; CURRENT FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: US 60/271,245
; PRIOR FILING DATE: 2001-02-15

; SEQ ID NO 1
; LENGTH: 67311
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-942-025-1

Query Match
Best Local Similarity 4.1%; Score 45; DB 10; Length 67311;
Matches 298; Conservative 0; Mismatches 360; Indels 7; Gaps 3;

QY 74 GCGTGAGGTGAAGGAGGGGCCACAGGCCAGTGGCGGTGTGCCACGAGCTGGTGCCTG 133
DB 29939 GCGCGGTGGCGCGCGCCAGACGACGCGTTCAGGACCTGGCGCGCGCGCTGTGGG 29998
QY 134 GACCCGAGTGTGTGTGATGCTGGCCCGCGGAGACCTACCGCTTCCTCGGTGGCAG 193
DB 29999 GCGTGTTCGCGCGCGCGAAGCGAGCACCCCGAAGCGCGGCTGGCTTGTGATGATGTGG 30058
QY 194 CTGTGGGCCCTGTGGGTGCTGGGAACCGGTTACCTGCCCGCCAGACAGTGGCGCTTCAG 253
DB 30059 GCACAGGAGCCGTGACGCTGGGCTGTGGCGCGGGCGCTGGCGAGCGGAGCCGG 30118
QY 254 AGCCACCGAAGCTGTGCTCCCGACCCCTCAGCCCTTGAGAGCGGCGAGGTGGCAGCTG 313
DB 30119 ACCTTGGCTGTGGCGGGCGCTGCGCTGGCGCGCGCTGTGGCGCACAGGCGGTAG 30178
QY 314 GTGAAGATGTCTCTGTGAGCTTGTGAGTGTGGTGGCTGAGCTGGTGAAGTTCATCTGGCACA 373
DB 30179 CGGAAGAGCTCACCCGAGCCCGAGCTGGACCTGCGGGCACGGTGTGGTGACCGGG 30238
QY 374 AGGGAATGAGCGCATCCAGCCCGGTGGGGTTCGAGGTGTCTCCAGAGGTGCGCAAC 433
DB 30239 GCACAGGGGAGCTGGGTGAGGCGGTGCGCGCGCACCTGTGTGGCGCGCGGGTGGCG 30298
QY 434 AGATGCTGTGTA--TCAGGGGCTTCACGGCAGAGAACAGGCGCGAGTACCACTGTGGCT 491
DB 30299 ACCTTGTGCTGAGCTGCGCGCGGGCTGGAGCGCCCGGGCGCGAGCTTGTGGAT 30358
QY 492 GCGTCAGGGCTCATCTGCGCTGGCTGCCACTTTCAGGTGGCACTGAGCCCGACCTC 551
DB 30359 CGCTCGCGAGTTCGCGCGCGAGACGGTGAACGTGTGGCGCGCTGCGACGTGTGCAAGC--G 30416
QY 552 TGTGATAGGGCCCTTCAGCCCGAGTTCGCCCGCGAGGCGCGCGAGGAGGTGACCTGCA 611
DB 30417 GAGAGAGTTCGCGCGTGTGCTGGCGCGCATCGACGCGCGCGCGCTGAGCGCGGTGT 30476
QY 612 CCTACTGTGGAGGCGCTGGCTCGGAACGTCGCGATGAGCCGTGAGCCCGACGCTGACTC 671
DB 30477 GCACCTGGCGCGCGCTTCGAGCAGCGGTGTCTGCCGCGCAGACGGCGCGCTCTC 30536
QY 672 CATTAGCGAGTTCAGAGGAGGACGGCGCTTCGAGCGCTTCGACAGGCGCTCCACAGGAGGAGGA 731
DB 30537 ---GCGGGTGTGGCGCGCGAAGGTGACGGGGCGCTGCACCTGCACGAGCTGACGCGGA 30593
QY 732 GGTGG 736
DB 30594 GCTGG 30598

RESULT 14
US-10-437-963-64043
; Sequence 64043, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
```


; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Bing
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 64043
; LENGTH: 1469
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_65226C.1
US-10-437-963-64043

Query Match 4.0%; Score 44.4; DB 18; Length 1469;
Best Local Similarity 50.0%; Pred. No. 0.087;
Matches 111; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

QY 254 AGCCACCGAGCTGTGCTCCACGCTCCAGCCCTCAGCCCTCAGAGCGGAGGTGGCAGCTG 313
DB 313 ATCCGCGGACGCGTGCAGCGTCCGCATCCTCTCCTCATCCGCTGTCTAGGAAGATG 372
QY 314 GTAAGATGTCTCTCTGGAGCTTTGAGGTGGTGGCTGAGGCTGTGAGGTTCATCTGGCACA 373
DB 373 GCGAGTTGGCGAGCTTGTCCATGGCGAGCTCGAGTTGCTGTAGTGCCTACTGCCCC 432
QY 374 AGGAAATGAGGCGCATCCAGCCCGTGGCGGTTCGAGGTGTCTCCAGGGTCCGGCAAC 433
DB 433 ATGAAGCTCCGCGAGCCATGGCGCGCGGCGGAGGGTGTATCATCCGCGGCGTGC 492
QY 434 AGATGCTGTGATCAAGGCTTCACGGCAGAGACCGGGCG 475
DB 493 TCTGCGACGACGCGCGCTCCGAGGAGGAGACGCTCG 534

RESULT 15
US-10-156-761-6923
; Sequence 6923, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 6923
; LENGTH: 4698
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(4698)
US-10-156-761-6923

Query Match 4.0%; Score 44.2; DB 15; Length 4698;
Best Local Similarity 43.4%; Pred. No. 0.098;
Matches 307; Conservative 0; Mismatches 393; Indels 7; Gaps 2;

QY 180 CTTCCGTGTGCAGCTGTGGGCGCTGTGGTGTCTGGGAAACGGTTTACCTGCCCCAGAC 239
DB 1119 CTACCGGCACGCTCTGTTGGCGCTCGCGAGCCGAGAAAGTCCGCGCGCGAGAACGG 1178
QY 240 AGTGGGCTTGAGAGCACCAGACCTGTGCTCCCGAGCCCTCAGCCCTGAGAGCGG 299
DB 1179 CGAGTGGCCGACGCGGACGCTGCACTCCGCTGGCAGGCTCCGAGGCGGTGTCTCG 1238
QY 300 GCAGTGGCAGCTGTGTAAGATGTCTCTCTGGAGCTTGAAGTGTGTGGTGTGAGGTGTGGA 359
DB 1239 GCACCGGCGCGCGCGCGCTCTCGCGCGGTGGCGCTCGCGATCGCGAGGCGGAGCG 1298
QY 360 GGTCAATCTGGCACAAGGAATGAGCGCATCCAGCCGCTGGGCGGTTCGAGGTGGTCTC 419
DB 1299 GGAGCGCGCGCGGCGCTCGCGCGCGCGCAAGCGCCCTCGACCTCTGCTACGTGCTCT 1358
QY 420 CCAGGCTCGGCAACAGATGTGTGATCAAGGGCTTCAAGGCAAGAACACAGGCGGAGTA 479
DB 1359 GCACGCTGCCCGGAGAGCGCGAGAACTCGCCCAACGAGGAGGAGGCGGTCCGCGAG 1418
QY 480 CCATGTGGCTGCTGCTGAGGCTCCATCTGCGCTGCGGCTGCGACCTTCCAGGTGGCACT 539
DB 1419 GCTCAGAGACGTGGCGGAGGCGGTACCGGGACTCGA CTTCCGCGCGCAACGAGGCGCA 1478
QY 540 GAGCCACGCTCTGTGATGAGGCGCTCAGCCCGAGCTTGCCCGCGGAGGCGAGCCAGGA 599
DB 1479 GCGC---GCCGCGAGCGAGCGCGGCGATCTGCGCCAGCGGCTCAGCGAGGTAGCGCAGA 1535
QY 600 GGGTGACTGCACTACTGTGGGAGGCGCTGGTTCGAAACGTTCGATGAGCCGTGAGCC 659
DB 1536 GACCGCGAGCGGTGCGCGCGGG---TGCGTGGACGACAGCGCGCGCGAGCGGACC 1591
QY 660 CACGCTGATCTTCATTAGCGAGCTGCCAGAGGAGGACGCGCGCTCGCAGCGCTGCCACA 719
DB 1592 CCGCACGCGCGGCGCTGGCGCGAGCGACCGCGAGAAAGACGCGCGCTCGCGCTGGGACA 1651
QY 720 GGAGGCGAGGAGGTGGCACTGTATCTCTCTGAAGGCTACTCCACGCGCGATGAGCTGGC 779
DB 1652 CCGCGCTGAGGCGTCCGCGCGGCGCTCCGAGCAGCGCCAGGAGGCGCGCTCCACCGAGT 1711
QY 780 CCGCACTGGAGATGCTGACCTCTCA CACACAGCTCTGATGATGATGATCCCGGGAGGAC 839
DB 1712 CCGCGCGGAGCTGACGCGCGCGCGCGCGCGGCGCGAGCGCGCGCGAGCGCGCGT 1771
QY 840 CCCTTCCCTGCTCACCTTACCTCAAGAGGCTGGGAGGCGCCAGGCGACCT 886
DB 1772 ACAGGCGGAGTTCCGTA CCGCGAGTCCCTGGCGGGGAGGAGCGT 1818

Search completed: March 21, 2005, 16:08:08
Job time : 701.694 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 16:50:48 ; Search time 1467.58 Seconds
(without alignments)
10400.615 Million cell updates/sec

Title: US-10-077-130-6_COPY_8500_8900
Perfect score: 401
Sequence: 1 gaagaccaggggtggcgcc.....caccagccagcaggagggg 401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsel1:*
9: gb_gsel2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	389.4	97.1	510	5	BX478907 DKFZp686N
2	326.6	81.4	354	2	BF353892 IL5-HT070
3	97.2	24.2	742	5	BU358283 603477905
4	79.6	19.9	895	2	BE731455 601567740
5	69.2	17.3	750	2	BE734558 601570867
6	64.8	16.2	591	5	BU439885 604144795
7	64	16.0	1060	9	CNS02279 AL221382 Tetraodon
8	62.2	15.5	429	4	BG383506 301607 MA
9	62.2	15.5	465	2	BF602347 267554 MA
10	60.8	15.2	3789	9	AY406292 Mus muscu
11	60.6	15.1	507	4	BM729096 UI-E-E01
12	60.6	15.1	895	6	CD244577 AGENCOURT
13	60.4	15.1	765	9	BX181691 Danio rer
14	59.8	14.9	371	2	BF824977 RC3-HN002
15	59.8	14.9	723	5	BQ446463 UI-H-EU1
16	59.2	14.8	553	7	CN302349 170005326
17	59	14.7	453	9	CG544711 OST142288
18	59	14.7	1192	8	CC247513 CH261-127
19	58.6	14.6	925	9	AL053013 Drosophila
20	58.4	14.6	1074	4	BJ730664 BJ730664
21	57.8	14.4	697	4	BI559709 603252581
22	57.8	14.4	801	4	BG912151 602812819
23	57.6	14.4	564	4	BG003227 MR3-GN018
24	57.4	14.3	747	7	CN039222 nm_31_f10

```
25 57.2 14.3 538 7 CN302346 170005326
26 57.2 14.3 588 5 BU436399 604145537
27 57 14.2 541 2 BE810116 218929 MA
28 56.6 14.1 505 2 BF087860 IL3-HT061
29 56.6 14.1 612 5 BU947303 1046C03.Y
30 56.4 14.1 698 7 CR787765 DKFZp468J
31 56.4 14.1 734 7 CO503162 GGEZCB102
32 56.4 14.1 782 5 BU438276 604146353
33 56.2 14.0 528 2 BE751691 203886 MA
34 56.2 14.0 537 4 BM666974 UI-B-CL1-
35 56.2 14.0 852 5 BX377135 BX377135
36 56 14.0 294 4 BM695897 UI-B-CL1-
37 56 14.0 437 4 BM703784 UI-B-CL1-
38 56 14.0 516 2 BF758009 OV3-CT055
39 56 14.0 525 2 BF087734 IL3-HT061
40 56 14.0 533 2 AW995362 QVO-BN004
41 56 14.0 541 6 CA416261 UI-H-FEO-
42 56 14.0 565 6 CA416975 UI-H-FEO-
43 56 14.0 575 5 BU072543 im44e07.Y
44 56 14.0 579 1 A1818399 wk78d03.X
45 56 14.0 579 6 CA414466 UI-H-EZ0-
```

ALIGNMENTS

```
RESULT 1
BX478907
LOCUS BX478907 510 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFZp686N21206_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
ACCESSION BX478907
VERSION BX478907.1 GI:31914176
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 510)
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., Oeanger, A.,
Fobo, G., Han, M. and Wiemann, S.
EST (Koehrer, K., Beyer, A., Mewes, H.W., Weil, B., Amid, C., et al.)
JOURNAL Unpublished (2003)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by BPFZ (Biomedical Research Center at the Heinrich-
Heine-University, Dueseldorf/Germany) within the cDNA sequencing
consortium of the German Genome Project. No sl sequence available.
This clone (DKFZp686N21206) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
Location/Qualifiers
source
1..510
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686N21206"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="686 (synonym: hlcc3)"
/notes="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"
```

```
ORIGIN
Query Match 97.1%; Score 389.4; DB 5; Length 510;
Best Local Similarity 99.7%; Pred. No. 2.4e-87;
Matches 390; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 11 GGGTGGGCCAGGGGAGGACGTGGAGCTGTGACCTGTACGGCGGGGAGCCCG 70
```


Db 6 GGTGGCGCCAGGGGAGGACGTGGAGCTGCTGTGAGCTGTCAAGGCGGGAACGCCCG 65
QY 71 TGCACTGGCTGAAGGACAGGAAGGCCATCCGAAGAGCCAGAAATGATGTTGGTGG 130
Db 66 TGCACTGGCTGAAGGACAGGAAGGCCATCCGAAGAGCCAGAAATGATGTTGGTGG 125
QY 131 AGGGACAGATGGCCATGCTGGTCACTCCGCGGGGCTCGCTCAAGGACGCGGCGAGTACA 190
Db 126 AGGGACAGATGGCCATGCTGGTCACTCCGCGGGGCTCGCTCAAGGACGCGGCGAGTACA 185
QY 191 CTGTGTGAGTGGAGGCTTCCAGAGGACAGCCAGCTCCATGTGGAAGAAAGCAAACT 250
Db 186 CTGTGTGAGTGGAGGCTTCCAGAGGACAGCCAGCTCCATGTGGAAGAAAGCAAACT 245
QY 251 GTTTCACAGAGGAGTGCACCAATCTGCAGGTGGAGGAGAAAGGCACAGCTGTGTTACGT 310
Db 246 GTTTCACAGAGGAGTGCACCAATCTGCAGGTGGAGGAGAAAGGCACAGCTGTGTTACGT 305
QY 311 GCAAGCGGAGACACCCCGGGCCACAGTACCTGCGCAAGGGCTCTTGGAGCTACGGG 370
Db 306 GCAAGCGGAGACACCCCGGGCCACAGTACCTGCGCAAGGGCTCTTGGAGCTACGGG 365
QY 371 CCTCAGGGAAGCACCAGCCGACCCAGGAGG 401
Db 366 CCTCAGGGAAGCACCAGCCGACCCAGGAGG 396

RESULT 2
BF353892
LOCUS IL5-HT0704-290600-108-e09 HT0704 Homo sapiens cDNA, mRNA sequence.
BF353892
DEFINITION EST.
VERSION BF353892.1 GI:11312966
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 354)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.P., Matsukuma,A., Baia,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&tl2=IL5-HT0704-
290600-108-e09&tl3=2000-06-29&tl4=1)
Seq primer: puc 18 forward
High quality sequence stop: 341.

FEATURES
source
1. .354
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0704"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI;

Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the puc 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

Query Match 81.4%; Score 326.6; DB 2; Length 354;
Best Local Similarity 98.8%; Pred. No. 1.5e-71;
Matches 329; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GAAGACCAAGTGGGTGGCGCCAGAGGAGACGTGGAGCTGCGCTGTGAGCTGTCAAGGCGG 60
Db 17 GAAGACCAAGTGGGTGGCGCCAGAGGAGACGTGGAGCTGCGCTGTGAGCTGTCAAGGCGG 76
QY 61 GGAACGCCGCTGCACTGGCTGAAGGACAGGAAGCCATCCGCAAGAGCCAGAAATGAT 120
Db 77 GGAACGCCGCTGCACTGGCTGAAGGACAGGAAGCCATCCGCAAGAGCCAGAAATGAT 136
QY 121 GTGCTCTGCGAGGCGACGATGGCCATGCTGGTCACTCCGCGGGGCTCGCTCAAGGACGCG 180
Db 137 GTGCTCTGCGAGGCGACGATGGCCATGCTGGTCACTCCGCGGGGCTCGCTCAAGGACGCG 196
QY 181 GCGAGTACACGTGTGAGGTGGAGGCTTCCAAGAGCACAGCCAGCTCCATGTGGAAGAA 240
Db 197 GCGAGTACACGTGTGAGGTGGAGGCTTCCAAGAGCACAGCCAGCTCCATGTGGAAGAA 256
QY 241 AAAGCAAACTGCTTCACAGAGGAGCTGACCAATCTGAGGTGGAGGAGAAAGGCACAGCT 300
Db 257 AAAGCAAACTGCTTCACAGAGGAGCTGACCAATCTGAGGTGGAGGAGAAAGGCACAGCT 316
QY 301 GTGTTACAGTGCAGAGCGGAGCACCCCGCGGCC 333
Db 317 TGTTGTCACGTGCAAGACGGAGCACCCCGCGGCC 349

RESULT 3

BF358283/c
LOCUS 742 bp mRNA linear EST 28-NOV-2002
DEFINITION 603477905F1 CSBQCHN71 Gallus gallus cDNA clone CHEST364j2 5', mRNA
sequence.
BF358283
ACCESSION BU358283
VERSION BU358283.1 GI:25866284
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 742)
AUTHORS Boardman,P.E., Sanz-Ezquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,
Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.
TITLE A Comprehensive Collection Of Chicken cDNAs
JOURNAL Curr. Biol. 12 (22), 1965-1969 (2002)
MEDLINE 22335534
PUBMED 12445392
COMMENT Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES
source

1. .742
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="White Leghorn, Hixex"
/db_xref="taxon:9031"
/clone="CHEST364j2"
/dev_stage="36"


```

/lab_host="DH10B"
/clone_lib="CSQCHN71"
/notes="Organ: hearts; Vector: pBluescript II KS(+);
Site 1: EcoRI; Site 2: NotI; This normalized library was
constructed from 1 million independent clones. cDNA
synthesis was initiated using an oligo(dT) primer, using
methylated C in the first strand synthesis reaction.
Following this first strand reaction, double-stranded cDNA
was blunt-ended, ligated to NotI adapters, digested with
EcoRI, size-selected, and cloned into the NotI and EcoRI
compatible sites of a custom modified MCS of the
pBluescript (KS+) vector. The library was normalized in 2
rounds using conditions adapted from Soares et al., PNAS
(1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6
(1996): 791, except that a significantly longer
reannealing hybridization was used."

```

ORIGIN

```

Query Match      24.2%; Score 97.2; DB 5; Length 742;
Best Local Similarity 54.9%; Pred. No. 8.7e-14;
Matches 192; Conservative 0; Mismatches 158; Indels 0; Gaps 0;

QY 1 GAAGACCACTGGTGGCGCCAGGGAGGACGTGGAGCTGCGCTGTGAGCTGTACAGGGCG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 GAAACAGGAGGTGGAGAGGGCGCACCGTTATGTTGCACTGTGAGCTGACGAAACCC 311
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 61 GGAACGCCGTGCTACTGGCTGAAGACAGAGAAAGGCATCCGCAAGAGCCAGAAATATGAT 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 310 AACGCACCCGTGGAGTGGAGGAAGAGATACGGTGTGCACTCAGGTGACAAATGATGAG 251
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 121 GTGGTCTGGAGGACAGTGCCTATGCTGTCATCCGGGGGCTCGCTCAAGACCGG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 250 GTCCGGCAAGAGGGACAGCTGTGAGCTCTTCATCTAGTATGCGAGGCTCAGGATGCG 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 181 GCGAGTACAGCTGTGAGTGGAGCTTCCAGAGCACAGCGAGCTTCATGTGGAAGAA 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 190 GGTGATTACAGTGTGACTCAGGGGATCAACAGACCATGTCATTCGAAGTCAAGTA 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 241 AAAGCAAACTGCTTCAAGAGAGCTGACCAATCTGCGAGTGGAGAGGAAAGCCAGCT 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 130 CTACCCGTGCTCTTTAAAGAGAGCTGAAAAATGTGGAGTCTGAAAGAGGGGAAACAGCT 71
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 301 GTGTTACGTGCAAGACGGAGCACCCCGCGGCACAGTGACCTGGCGAA 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 70 ATCTTGCACTGCGAGATTCCAAAGCCAGATGCTCCAGTTGAGTGGAAAAA 21
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 4

```

BE731455
LOCUS      601567740F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3842552 5',
DEFINITION mRNA sequence.
ACCESSION BE731455
VERSION    BE731455.1 GI:10145447
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 895)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW535 row: P column: 09
High quality sequence stop: 749.

```

FEATURES

source

```

Location/Qualifiers
1..895
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3842552"
/tissue_type="choriocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_21"
/notes="Organ: placenta; Vector: pOTB7; Site 1: XhoI;
Site 2: EcoRI; cDNA made by oligo-dT priming.
Directionally cloned into EcoRI/XhoI sites using the
following 5' adaptor: GGCACGAG(G). Size-selected >500bp
for average insert size 1.8kb. Library constructed by
Ling Hong in the laboratory of Gerald M. Rubin (University
of California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene) and Superscript II RT (Life Technologies)."
```

ORIGIN

```

Query Match      19.9%; Score 79.6; DB 2; Length 895;
Best Local Similarity 54.0%; Pred. No. 2.4e-09;
Matches 163; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 30 CGTGAGCTGCGCTGTGAGCTGTACGGCGGGACGCCGCTGCACTGGCTGAAGACAG 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 CGGTGTGTGAGTGTGAGTGTCCGGGGCCAGCCCGAGGTGCGGTGGTTCAGGGCAG 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 90 GAAGGCCATCCGCAAGAGCCAGAAAGTATGTGTCTGCGAGGGCCAGATGCCATGCT 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 231 TCAGGAGCTGCAGCCCGGCCCAAGTACGAGCTGTGTGATGATGCTCTTACCGCAAGCT 290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 150 GGTCTATCCGGGGGCTCGCTCAAGGACGGGGGAGTACACGTGTGAGTGGAGGCTTC 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 291 GATCATCATGTATGTCTCACGAGAGGAGGACACCTACACCTGTGACGCGCGGTGT 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 210 CAAGAGCACAGCAGCGCTCCATGTGGAAGAAAAAGCAAACTGCTTCACAGAGAGCTGAC 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 CAAGACCACTGCACAGTCTTCGTGGAAAGAGCAATCCATCACCATTGCGGGGTCTGCA 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 270 CAATCTGAGCTGGAGAGAAAGGCACAGCTGTGTTCAGTGTGCAAGACGGAGCACCCCGC 329
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 GGACGTGACAGTGTGAGAGCCGCTCTGCTGCTGTTGAGTGTGAGACCTCCATCCCCCTC 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 330 GG 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 471 AG 472
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 5

```

BE734558
LOCUS      601570867F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3845393 5',
DEFINITION mRNA sequence.
ACCESSION BE734558
VERSION    BE734558.1 GI:10148550
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 750)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Ling Hong/Rubin Laboratory
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: LLCW543 row: f column: 18

```


REFERENCE 1
 AUTHORS Roset Crollius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fzames, C., Wincker, P., Brotier, P., Quetier, F., Saurin, W., and Weissenbach, J.
 TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence
 JOURNAL Nat. Genet. 25 (2), 235-238 (2000)
 MEDLINE 20296633
 PUBMED 10835645
 REFERENCE 2
 AUTHORS Roset Crollius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fzames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A., and Weissenbach, J.
 TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis
 JOURNAL Genome Res. 10 (7), 939-949 (2000)
 MEDLINE 20359837
 PUBMED 10899143
 REFERENCE 3
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
 COMMENT - Web : www.genoscope.cns.fr
 This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.
 FEATURES
 source
 1..1060
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="183C20"
 /clone_lib="G"
 /note="Genoscope sequence ID : COAG183BB10SP1-end : PUC-ori"
 ORIGIN
 Query Match 16.0%; Score 64; DB 9; Length 1060;
 Best Local Similarity 55.7%; Pred. No. 2,1e-05;
 Matches 118; Conservative 2; Mismatches 92; Indels 0; Gaps 0;
 QY 12 GGTGGCCGAGGGAGGAGCTGGAGCTGGCTGGTGTGAGCTGTGACGGGGGAGACGCCCGT 71
 Db GGCACTCCCGCGGAGGAGCGAGAGCTGAGCTGTGAGATCAACACAGACAGCTCAGCAT 893
 QY 72 GCACCTGGCTGAAGGACAGAGAGCCATCGCAAGAGCCAGAGTATGATGTTGCTGCGA 131
 Db TCGTTGGCTGAAGRACGGCCGTTTGTATCAGGACAGAGCCCAAGTATGATGAGCTTAT 833
 QY 132 GGCACAGATGGCCATGCTGGTTCATCCGGGGGCTCGCTCAAGGACGGGGCGAGTACAC 191
 Db GAAGAACCTGGCAAGACTGGTGTATCAGGACAGAGCCACCATCAGAGACTCTGGGATACTG 773
 QY 192 GTGTGAGTGGAGGCTTCCAGAGACAGACCA 223
 Db CTGCGAGGACAGCGGTGTGCTCGGAGCCA 741
 RESULT 8
 LOCUS BG383506 429 bp mRNA linear EST 12-MAR-2001
 DEFINITION 301607 MARC 1P1G Sus scrofa cdna 5', mRNA sequence.
 ACCESSION BG383506
 VERSION BG383506.1 GI:13307978
 KEYWORDS EST.
 SOURCE Sus scrofa (pig)
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE 1 (bases 1 to 429)
 AUTHORS Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,

Vallet, J., Wise, T., Rohrer, G.A., Pertea, G., Sultana, R., Quackenbush, J., and Keefe, J.W.
 TITLE Porcine gene discovery by normalized cDNA-library sequencing and EST cluster assembly
 JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
 MEDLINE 22213789
 PUBMED 12226715
 COMMENT Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCACGTCACGAGC
 Plate: 87 row: O column: 10
 Seq primer: ATTAGGTGACACTATAG.
 Location/Qualifiers
 1..429
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_lib="MARC 1P1G"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from day 11, 13, 15, 20, and 30 embryos."
 ORIGIN
 Query Match 15.5%; Score 62.2; DB 4; Length 429;
 Best Local Similarity 55.3%; Pred. No. 5,2e-05;
 Matches 121; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
 QY 10 TGGGTGGCGCCAGGGAGGACGTGGAGCTGGCTGTGAGCTGTACGGCGGGAAACGCC 69
 Db TGTGTGGCCCGGGGATCCGCTGGTGTGAGCTGTGAACCTGCAGGGCTGGCGCCCTT 234
 QY 70 GTGCACCTGGCTGAAGGACAGGAGCCATCCGACAGAGCCAGAGTATGATGTGCTGTC 129
 Db GTCTTTTGGAGCCCAACGGGAGGCCAGTTTCAGAGGGCGAGGCTCGAGCTCCATGCT 294
 QY 130 GAGGCGCCCGCCCGCATCTGCTATCCGGGGGCTCGCTCAAGGACCGGGCGAGTAC 189
 Db GAGGCGCCCGCCCGCATCTCTGTCATCGGGCTGCAGACCCGGCCCATCGGGCCTCTAC 354
 QY 190 ACGTGTGAGTGGAGGCTTCCAAAGAGCACAGCCAGCCTC 228
 Db ACCTGCCAGTCCGGGAGCAGCCCGGGGCTCCAGCCTC 393
 RESULT 9
 LOCUS BF602347 465 bp mRNA linear EST 25-APR-2001
 DEFINITION 267554 MARC 3BOV Bos taurus cdna 5', mRNA sequence.
 ACCESSION BF602347
 VERSION BF602347.1 GI:11699571
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
 REFERENCE 1 (bases 1 to 465)
 AUTHORS Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.M., Rohrer, G.A., Chitko-McKown, C.G., Pertea, G., Holt, I., Karamycheva, S., Liang, F., Quackenbush, J., and Keefe, J.W.

TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
JOURNAL	Genome Res. 11 (4), 626-630 (2001)
MEDLINE	21180013
PUBMED	11282978
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single passes sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.

```

seq primer: alltaggagacaciatag.
FEATURES
    source
        1..465
            Location/Qualifiers
                organism="Bos taurus"
                mol_type="mRNA"
                db_xref="taxon:9913"
                tissue_type="pooled"
                lab_hosts="DH108"
                clone_lib="MARC 380V"
                notes="vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
                Library made from pooled tissue from marrow, alveolar
                macrophage, ovary, fetal semitendinosus muscle, and fetal
                longissimus muscle."
ORIGIN
    Query Match          15.5%; Score 62.2; DB 2; Length 465;
    Match Local Similarity 55.3%; Pred. No.5.3e-05;
    Matches 121; Conservative 0; Mismatches 98; Indels 0; Gaps 0;
    10 TGGGTGGCGCCAGGGAGGACGTGGAGCTGCGCTGTGAGCTGTGCACGGCGGGAGCGCCC 69
    83 TGGTGGCCCCCGGGGAGCCAGTGGTGTGTGAGCTGTGACTGTCTCAGGGCTGTGTGCCCTG 142
    70 GTGCACCTGGCTGAAGGACAGGAGGCGCATCCGCCAAGAGCCAGAGATATGATGTGGTGTGC 129
    143 GTCTTCTTGGAGCCATAACGGGGAAGCGGTGCAACAGCGGAGGGCTGTGAGCTCCGAGCC 202
    130 GAGGGCACCAGTATGCCATGTGTGTCTCGCGGGGCTCGCTCAAGGACGCGGGCGAGTAC 189
    203 GAAGGACCCGCGCGTCTCTTGTGATCCGGGCTGAGACCTTGGCCCCAGCGGCCTCTTAC 262
    190 ACTGTGAGGTGGAGCGTTTCCAAAGACACAGCCAGCCTTC 228
    263 ACCTGCCAGTGTGGGCGACACCGGGGGCCCCCAGCCTTC 301

```

ORIGIN		longissimus muscle.	
Query Match	15.5%	Score 62.2;	DB 2; Length 465;
Best Local Similarity	55.3%;	Prod. NO. 5.3e-05;	
Matches 121;	Conservative	0; Mismatches 98;	Indels 0; Gaps 0;
QY	10	TGGTGTGGCGCCAGGGAGGACGTGGAGCTGCCTCTGAGCTGTACACGGGCGGGNACGGCC	69
Ddb	83	TGCGTGGCCCCCGGGAGGCCAGTGTGTCTGAGCTGTGAACCTGTCCAGGGCTGGTGCCCTG	142
QY	70	GTGCATCTGGCTGAAGGACAGGAGGCCATCCCGAAGAGCCAGAAAGTATGATGTGGTCTGC	129
Ddb	143	GTCTTCTGGAGCCATAACCGGGAAGCCGGTGTCAAACAGCGCAGGGCCTGGAGCTCCGAGCC	202
QY	130	GAGGGACAGATGSCCATGTGTGTCTATCCCGGGGCTCGCTCAGAGAGCGGGCGGAGTAC	189
Ddb	203	GAAGACCCCGCGCTCTCTTGATCCGGGCTGCAGACCTTGGCCCCACGAGGCCTCTAC	262
QY	190	ACGTGTGAGGTGGAGGCTTCCAAGAGCAAGCAGCGCTTC	228
Ddb	263	ACCTGCCAGTGTGGGCGAGCACCGGGGGCCCCCAGCGCTTC	301

RESULT 10	AY406292	3789 bp	DNA	linear	GSS 15-DEC-2003				
LOCUS	AY406292								
DEFINITION	Mus musculus MYBPC3 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.								
ACCESSION	AY406292								
VERSION	AY406292.1	GI:39762266							
KEYWORDS	GSS.								
ORGANISM	Mus musculus	Mus musculus (house mouse)							
	Mus musculus								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
	1 (bases 1 to 3789)								
AUTHORS	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejaritwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sainsky, J.J., Adams, M.D. and Cargill, M.								
TITLE	Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios								

JOURNAL	Science 302 (5652), 1960-1963 (2003)
PUBMED	14671302
REFERENCE	2 (bases 1 to 3789)
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.
TITLE	Direct Submission
JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES	Location/Qualifiers
source	1..3789 /organism="Mus musculus" /mol_type="genomic DNA" /db_xref="taxon:10090" <1..>3789 /gene="MYBPC3" /locus_tag="HMC2508"
gene	
ORIGIN	
Query Match	15.2%; Score 60.8; DB 9; Length 3789;
Best Local Similarity	49.4%; Pred. No. 0.00016;
Matches 158; Conservative	0; Mismatches 162; Indels 0; Gaps 0;
Qy	46 GAGCTGTCACGGCGGGGAAGCCCGTGCACCTGTGCTGAAGGACAGGAGGCATCCGCAAG 105
Db	1134 GAACCTGGCAGATCGGAGCGCTGAAGTCAAGTGGCTTAAGATCGACAGGAGATCCAGATG 1193
Qy	106 AGCAGAGATGATGTGTGTCCTCGAGGGGACGATGGCCATGTGTGTCATCCCGGGGCC 165
Db	1194 AGTGGCAGGTACATCTTCGAGTCCGTCCGTGGTCCCAAGCGCACCCCTGCACCATCAGCCAGTGC 1253
Qy	166 TCCGCTCAAGGACCGCGCGAGTACAGTGTGAGGTGGAGGCTTCCAAGAGCACAGCCAGC 225
Db	1254 TCACCTGGCTGACGACGACGACGCTACCAAGTGTGTGTGGGGGGCGAAGTGTGACACGGAG 1313
Qy	226 CTCATGTGAGAGAAAAGCAAACTCCTTCACAGAGGAGCTGACCAATCTGCAGGTGGAG 285
Db	1314 CTCCTTTGTCAAGAGAGCCCGGGTGTGATCACTCGGTCCCTGGAGACCCAGCTGGTGTGATG 1373
Qy	286 GAGAAAAGGCACAGCTGTGTTCACGTGCAAGACGGAGCACCCCGCGGCCACAGTGACCTGG 345
Db	1374 GTGGGTACGGGTGGAGTTTGAGTGTGAGGTCTCAGAGAGAGGGGCCCCAGTCAATGG 1433
Qy	346 CGCAAGGGCCTCTTGGAGCT 365
Db	1434 CTGAAGGATGGGGTTTGAGCT 1453
RESULT 11	
BM729096	
LOCUS	BM729096 507 bp mRNA linear EST 01-MAR-2002
DEFINITION	UI-E-E01-aiw-i-04-0-UI.r1 UI-E-E01 Homo sapiens cDNA clone
ACCESSION	UI-E-E01-aiw-i-04-0-UI 5', mRNA sequence.
VERSION	BM729096
Accession	BM729096.1 GI:19050429

LOCUS	BM729096	507 bp	mRNA	linear	EST 01-MAR-2002
DEFINITION	UI-E-EOI-aiw-i-04-0-UI.r1 UI-E-EOI Homo sapiens cDNA clone UI-E-EOI-aiw-i-04-0-UI 5', mRNA sequence.				
ACCESSION	BM729096				
VERSION	BM729096.1				
KEYWORDS	EST.				
SOURCE	GI:19050429				
ORGANISM	Homo sapiens (human)				
REFERENCE AUTHORS TITLE	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
	1 (bases 1 to 507)				
	Bonaldo,M.F., Lennon,G. and Soares,M.B.				
	Normalization and subtraction: two approaches to facilitate gene discovery				
JOURNAL	Genome Res. 6 (9), 791-806 (1996)				
MEDLINE	97044477				
PUBMED	8889548				
COMMENT	Contact: Soares, MB				
	Coordinated Laboratory for Computational Genomics University of Iowa				

375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565

Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research Genetics (www.reagen.com).
Seq primer: M13 Reverse.

FEATURES
source
1. 507
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EQ1-aiw-i-04-0-UI"
/tissue_type="fetal eye"
/dev_stages="fetal"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-E-EQ1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; UI-E-EQ1 is a normalized cDNA library containing the following tissue(s): fetal eye. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is GCGGTATACC. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

ORIGIN

Query Match 15.1%; Score 60.6; DB 4; Length 507;
Best Local Similarity 54.8%; Pred. No. 0.00013;
Matches 120; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 10 TGGGTGGCGCCAGGGAGACGTGAGCTGCCTGTGAGCTGTCAAGAGCGGGGAGCGCC 69
DB 189 TGTGTGGCCCCCGGGGAGCCAGTGTGTGAGCTGTGAACCTCTCCCGGCTGCGCGCC 248

QY 70 GTGCACCTGCTGAAGCAGGAGCCATCCGACAGAGCCAGAGTATGATGTGCTCTGC 129
DB 249 GTGTCTGAGCCCAATATGGAGGCCCGTGACGAGGGCGAGGGCTTAGAGCTCCATGCC 308

QY 130 GAGGCGACGATGGCCATGCTGTCTATCCGCGGGGCTCGCTCAAGAGCGGGGCGAGTAC 189
DB 309 GAGGCGCCCCCGCGAGTCTCTGTCATCAGGCTGCAGGCCAGCCCATGCGAGGGCTCTAC 368

QY 190 ACGTGTGAGTGGAGCTTCCAAAGCAGCAGCAGCCCTC 228
DB 369 ACCTGCCAGTCTGGAGAGCGCCCCCGAGGCCCAAGCCCTC 407

RESULT 12

CD244577
LOCUS CD244577 895 bp mRNA linear EST 22-MAY-2003
DEFINITION AGSCOURT 14099068 NIH MGC 180 Homo sapiens cDNA clone IMAGE:30378371 5', mRNA sequence.
ACCESSION CD244577
VERSION CD244577.1 GI:31005041
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 895)

AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: NDAM444 row: b column: 12
High quality sequence start: 22
High quality sequence stop: 631.

FEATURES

source

1. 895
/location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:30378371"
/lab_host="DH10B-Ton A (T1 and T5 phage resistances)"
/clone_lib="NIH MGC 180"
/note="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV (destroyed); Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.68 kb. Library was constructed by (Invitrogen). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 15.1%; Score 60.6; DB 6; Length 895;
Best Local Similarity 54.8%; Pred. No. 0.00015;
Matches 120; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 10 TGGGTGGCGCCAGGGAGACGTGAGCTGCCTGTGAGCTGTCAAGAGCGGGGAGCGCC 69
DB 60 TGTGTGGCCCCCGGGGAGCCAGTGTGTGAGCTGTGAACCTCTCCCGGCTGCGCGCC 119

QY 70 GTGCACCTGCTGAAGCAGGAGCCATCCGACAGAGCCAGAGTATGATGTGCTCTGC 129
DB 120 GTGTCTGAGCCCAATATGGAGGCCCGTGACGAGGGCGAGGGCTTAGAGCTCCATGCC 179

QY 130 GAGGCGACGATGGCCATGCTGTCTATCCGCGGGGCTCGCTCAAGAGCGGGGCGAGTAC 189
DB 180 GAGGCGCCCCCGCGAGTCTCTGTCATCAGGCTGCAGGCCAGCCCATGCGAGGGCTCTAC 239

QY 190 ACGTGTGAGTGGAGCTTCCAAAGCAGCAGCAGCCCTC 228
DB 240 ACCTGCCAGTCTGAGAGAGCCCCCGAGGCCCAAGCCCTC 278

RESULT 13

CD244577
LOCUS BX181691 765 bp DNA linear GSS 13-MAR-2003
DEFINITION Danio rerio genomic clone DKEY-190D1, genomic survey sequence.
ACCESSION BX181691
VERSION BX181691.1 GI:28013507
KEYWORDS GSS.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 765)
Humphray, S.J., Huckle, E. and Durham, J.L.
Direct Submission
Submitted (13-MAR-2003) The Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Unpublished
This sequence was generated from the SP6 end of BAC 190D1. 190D1 is part of the Daniokey BAC Library created by R. Plasterk and N.V. Keygene. Further details:

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: March 20, 2005, 16:50:48 ; Search time 3663.47 Seconds
(without alignment)
10400.615 Million cell updates/sec

Title: US-10-077-130-6_COPY_2000_3000
Perfect score: 1001
Sequence: 1 cgctcaccatcgaggagtg.....ctgaggcaggggccaatgcc 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_est8:*

9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
C 1	654	65.3	740	CA430292	UI-H-FL1-
C 2	286.4	28.6	553	CN302349	CN302349 170005326
C 3	281.6	28.1	632	CF180572	CF180572 815696 MA
C 4	262	26.2	538	CN302346	CN302346 170005326
C 5	255.6	25.5	522	AQ595824	AQ595824 HS_2132_B
C 6	243.2	24.3	621	CF179218	CF179218 81316 MA
C 7	217.6	21.7	1142	BG680164	BG680164 60282628
C 8	217	21.7	401	AW898791	AW898791 CM0-NN007
C 9	216.4	21.6	260	BQ365446	BQ365446 RC6-GN007
C 10	203.4	20.3	658	AZ383233	AZ383233 1M0140117
C 11	202.4	20.2	574	A1595154	A1595154 ml39h10_Y
C 12	166.8	16.7	631	BG793386	BG793386 UTSW_SM11
C 13	153.2	15.3	591	BH439885	BH439885 604144795
C 14	149.4	14.9	343	BE159678	BE159678 MR0-HT040
C 15	146.6	14.6	397	BZ837763	BZ837763 CH240_248
C 16	142.2	14.2	373	AA061264	AA061264 ml39h10_r
C 17	136.6	13.6	786	CD099595	CD099595 AGENCOURT
C 18	134.6	13.4	236	CV426951	CV426951 RC6-EN008
C 19	123.4	12.3	588	BH439885	BH439885 604144795
C 20	122	12.2	472	CE495069	CE495069 tigr-g88-
C 21	113.8	11.4	1192	CC247513	CC247513 CH261-127
C 22	98.4	9.8	655	AU119815	AU119815 AU119815
C 23	98.4	9.8	724	AU141131	AU141131 AU141131
C 24	96	9.6	508	CD674219	CD674219 fs08a05.Y

25	94.4	9.4	770	4	BG819323	BG819323 602781512
26	93	9.3	4130	3	AK048144	AK048144 Mus muscu
27	93	9.3	5558	3	AK029863	AK029863 Mus muscu
28	92	9.2	179	4	BG989628	BG989628 PM2-HT117
29	91.2	9.1	569	4	BM254228	BM254228 S15592 MA
30	88.4	8.8	2066	9	AY399870	AY399870 Homo sapi
31	88.4	8.8	2274	3	CR590353	CR590353 full-length
32	81.4	8.1	886	5	BUS38739	BUS38739 AGENCOURT
33	81.2	8.1	205	7	CV419760	CV419760 RC4-CN018
34	80.4	8.0	612	5	BU947303	BU947303 1046C03.Y
35	79.8	8.0	575	5	BU072543	BU072543 1m4607.Y
36	79.8	8.0	886	5	CR607106	CR607106 full-length
37	79.8	8.0	886	5	BX377136	BX377136 BX377136
38	79.8	8.0	1085	4	BM546177	BM546177 AGENCOURT
C 39	73.4	7.9	980	5	BX434387	BX434387 BX434387
C 40	78.4	7.8	796	5	BX381923	BX381923 BX381923
41	77.4	7.7	2066	9	AY399872	AY399872 Mus muscu
42	76.4	7.6	1098	4	BM543727	BM543727 AGENCOURT
43	74.8	7.5	801	4	BG912151	BG912151 602812819
44	74.6	7.5	724	5	BP144357	BP144357 BP144357
C 45	74.2	7.4	533	2	AW995362	AW995362 QV0-BN004

ALIGNMENTS

RESULT 1
CA430292/c
LOCUS CA430292 740 bp mRNA linear EST 07-NOV-2002
DEFINITION UI-H-FL1-bf1-i-01-0-UI-s1 NCI CGAP FL1 Homo sapiens cDNA clone
UI-H-FL1-bf1-i-01-0-UI 3', mRNA sequence.
ACCESSION CA430292
VERSION CA430292.1 GI:24793018
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Cartrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 740)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Martin
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bentos-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=Yes.

FEATURES

Location/Qualifiers
1..740
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="UI-H-FL1-bf1-i-01-0-UI"
/tissue_type="Cell lines"
/dev_stages="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP FL1"
/note="Organ: Chondrosarcoma; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site 1: Ecor I; Site 2: Not I; NCI CGAP FL1 is a normalized cDNA library derived from a pool of mRNA obtained from 4 cell lines from grade III chondrosarcoma tissues. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an Ecor I adaptor, digested with Not I, and cloned directionally."

into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dfr)18 tail. The sequence tag for this library is GAGGTCGGTG. The cell lines were provided by Dr. James Martin from the University of Iowa.
TAG_TISSUE=Human Chondrosarcoma Grade 3 cell line mix
TAG_LIB=UT-H-FL1
TAG_SEQ=GAGGTCGGTG"

ORIGIN

Query Match 65.3%; Score 654; DB 6; Length 740;
Best Local Similarity 100.0%; Pred. No. 5e-128;
Matches 654; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCTCACATCCGGAGGTGCGCCAGCCTGACGCGGCGCAGCTGAAGTTCGTGGCCA 60
Db 695 CGCTCACATCCGGAGGTGCCGCCAGCCTGACGCGGCGCAGCTGAAGTTCGTGGCCA 636

Qy 61 ACGCCATTGAGAGCAGCATCCGATGAGGTTCGCGGCGGCCCCAGGCGTGAAGTCCCAACA 120
Db 635 ACGCCATTGAGAGCAGCATCCGATGAGGTTCGCGGCGGCCCCAGGCGTGAAGTCCCAACA 576

Qy 121 AGCCGCCAGCGCAGCTGCCCGGAGGTGCTGGCTCGGCTGCAGAGAGCGCAGCTGC 180
Db 575 AGCCGCCAGCGCAGCTGCCCGGAGGTGCTGGCTCGGCTGCAGAGAGCGCAGCTGC 516

Qy 181 TGGCTGAGCTGTGATGATCAGCTCCGCTGCTGAGCTGGCTGAAGGATGCTGCACACTGT 240
Db 515 TGGCTGAGCTGTGATGATCAGCTCCGCTGCTGAGCTGGCTGAAGGATGCTGCACACTGT 456

Qy 241 CCCAGAGCCCAAGTATGAGGTGAGCATCGGCGGCGGCGGCGGCGGCTCTCTTTGTCGAG 300
Db 455 CCCAGAGCCCAAGTATGAGGTGAGCATCGGCGGCGGCGGCGGCGGCTCTCTTTGTCGAG 396

Qy 301 ATGTGGCCCGGAGCATGTCAGGCTCTACGATGCTGTCAGCGCGGCGGCGGCGGCTCGCCT 360
Db 395 ATGTGGCCCGGAGCATGTCAGGCTCTACGATGCTGTCAGCGCGGCGGCGGCGGCTCGCCT 336

Qy 361 ACCAGCTCTCCGTGCAAGGCTCGCGGCTTTCTGCACAAAGGACATGCGCGGCGGCGGCTGTG 420
Db 335 ACCAGCTCTCCGTGCAAGGCTCGCGGCTTTCTGCACAAAGGACATGCGCGGCGGCGGCTGTG 276

Qy 421 TGGATGCCGTGGTGGGGGCCCGCGCATGTTGATGTGAGACTCCGAAAGCCCACTCC 480
Db 275 TGGATGCCGTGGTGGGGGCCCGCGCATGTTGATGTGAGACTCCGAAAGCCCACTCC 216

Qy 481 ACGTGCACTGTTACAGGATGGATGGAGCTGGGCGCACTCCGCTGAGCGCTTCTTTCAGG 540
Db 215 ACGTGCACTGTTACAGGATGGATGGAGCTGGGCGCACTCCGCTGAGCGCTTCTTTCAGG 156

Qy 541 AGGATGTGGGAGCGCGCAGCCGCTGCTGGCAGGCACAGTCAACAGGAGGATGAAGGCA 600
Db 155 AGGATGTGGGAGCGCGCAGCCGCTGCTGGCAGGCACAGTCAACAGGAGGATGAAGGCA 96

Qy 601 CCTACTCCTCGCGTGGGAGGAGCTCTGTGACTTCGCGCTCCGCTCTCTG 654
Db 95 CCTACTCCTCGCGTGGGAGGAGCTCTGTGACTTCGCGCTCCGCTCTCTG 42

RESULT 2

CN302349
LOCUS 17000532627462 GRN_ES Homo sapiens cDNA 5', mRNA sequence. EST 16-MAY-2004
DEFINITION CN302349
ACCESSION CN302349.1 GI:47318763
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 553)
Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Fisk,G.J.,

Li,Y., Xu,C., Fang,R., Guegler,K., Rao,M.S., Mandalam,R.,
Lebkowski,J and Scanton,L.W.
Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 553 Std Error: 0.00.

FEATURES

Location/Qualifiers
source
1..553
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_lib="GRN_ES"
/notes="oligo dT primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Query Match 28.6%; Score 286.4; DB 7; Length 553;
Best Local Similarity 78.2%; Pred. No. 3.1e-50;
Matches 344; Conservative 0; Mismatches 96; Indels 0; Gaps 0;

Qy 562 GCGTGGTGGCAGCCACAGTCCAGGAGGATGAAGGACCTACTCTCGCGGTGGCG 621
Db 4 GCGTGGTAGTGCAGAGTGGGCAAGCAGATGCTGGGGAGTACAGCTCGAGGCTGGG 63

Qy 622 AGGACTCTTGGAGCTTCGCGCTCTCTGAGCCCAAGGTGGTGTGCTAAGGAGC 681
Db 64 GCAGAGAGTCTCTTTCACTGCATCACAGAGCCAGGAGTGTGTCGAAGGAGC 123

Qy 682 AGCTGGCAGCCAGGAAGTGCAGGAGAGAGCAGGAGCCAGTGCACAGTGCAGG 741
Db 124 AGTTGGTGCAATAGAGTGGGACTGAGGAGGCGGCGGCGGCGGCGGCGG 183

Qy 742 TGGCCAGGCGCCAGCGAGTGCAGTGTGACAGGATGGAGAGTGCAGCTCAGCT 801
Db 184 TGGCCAGGCGCCAGCGAGTGCAGTGTGACAGGATGGAGAGTGCAGCTCAGCT 243

Qy 802 CGAAAGTGTGCAGGAGCCACAGGCTGCAGGCGGCGGCTGGTGTGTCAGCAGGCGC 861
Db 244 CGAAAGTGTGCAGGAGCCACAGGCTGCAGGCGGCGGCTGGTGTGTCAGCAGGCGC 303

Qy 862 AGCGGATGCCGGGAGTATAGCTGCAGAGCTGGGGCCAGCGGCTCTCTTCATCTGG 921
Db 304 AGCGGATGCCGGGAGTATAGCTGCAGAGCTGGGGCCAGCGGCTCTCTTCACCTGG 363

Qy 922 ATGTCGAAGAGCCCAAGTGTGTTTGCAGAGCAGGAGTGGGCGGCGGCGGCGG 981
Db 364 ATGTTTCAGAGCCCAAGCGGTGTTTGCAGAGGAGTGGGCGGCGGCGGCGG 423

Qy 982 CTGAGGAGGCGGCGCAATGCC 1001
Db 424 CCGAGGCGGCGGCGCAATGCC 443

RESULT 3

CF180572/c
LOCUS CF180572 632 bp mRNA linear EST 28-JUL-2003
DEFINITION 815696 MARC 3P1G Sus scrofa cDNA 3', mRNA sequence.
ACCESSION CF180572
VERSION CF180572.1 GI:33292348
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 632)
Smith, T.P.L., Freking, B.A., Ford, J.J., Vaillet, J.L., Fox, J.,
Wise, T.A., Nonneman, D.J., Wray, J.E. and Keele, J.W.
A second set of porcine ESTs from a pooled-tissue normalized
library
Unpublished (2003)
Contact: Smith, TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: SRG8009 row: O column: 4
Seq primer: TAGAAGGCACAGTCAGG.

FEATURES
source
1..632
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 3P1G"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including brain, liver, muscle, placenta/endometrium,
ovary, testes, and bone marrow."

ORIGIN
Query Match 28.1%; Score 281.6; DB 7; Length 632;
Best Local Similarity 68.4%; Pred. No. 3.3e-49;
Matches 405; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

Qy 406 TGGGGGCGAGCTGTGTGATGCGGTGGGGGCGGGCGGCGAGTTTGTAGTGTGAGACCT 465
Db
591 TGGCTCGAGTGAAGTGAAGCATGGCAGGGGCCAGTGCACGATGAGCTGTGAGTGG 532
Qy 466 CCGAAGCCACGCTCCACGTGCACTGGTACAAAGGATGGATGGATGGGCGGCGACCTCCGGTG 525
Db 531 CCCAGGCCAGAGGAGGTGACTTGGTACAAAGGATGGGAAGCTGAGTTCGAGCTCAA 472
Qy 526 AGCGTCTTCGAGAGAGATGGGGAGCGGGCAACGGCTGTGTCACGCCACAGTCACCA 585
Db 471 AAGTCGCTGTGGAGGCCAAGGGCTGCAACCCGGA---GGCTGTGTGTGTCAGCAAGTGGGCA 415
Qy 586 GGCAGGATGAAGGCACCTACTCTCCGCGTGGGCGAGGACTCTGTGGACTTCCGGCTCC 645
Db 414 AGGCAGATGATGGGGAGTACAGCTGCGAGGCTGGGGGTGAGAAGGTCTCTTCCACATGG 355
Qy 646 GCGTCTCTGAGCCCAAGGTGGTGTGTCTAAGGAGCAGCTGGCAGCAGGAGCTGTCAGG 705
Db 354 ATGTCACAGAGCCCTCAGTGTGTGTGTCAGGAGCAGCAGCAGCAGTGAAGTGCAG 295
Qy 706 CAGAGCAGGAGCCAGTCCACATGAGCTGCGAGGTGCGCCAGGCCAGACGGAGGTGA 765
Db 294 CCATGGCGGAGCCAGTGCACATGAGCTGTGAGGTGGCCCGAGCCCGACAGAGGTGA 235
Qy 766 CTGGTACAGGATGGGAAGTGAAGTCCAGCTCCAGTCCAAAGTGTGCATGAGGCCACAG 825
Db 234 CATGGTACAAAGATGGGAAGAGCTGACTTCAAGCTCAAAAAGTGTGTGGAGGCCCCCC 175
Qy 826 GCTGACCGCAGGCTGGTGTGTGACAGCAGGAGGCGAGCGGATCCGGGGAGTATAGCT 885
Db 174 CTTGCAACCGAGGCTGTGTGTGTGTCAGCAGCGGGCAAGGCAGATGATGGGNGTACAGCT 115
Qy 886 CGCAGGCTGGGGGCCAGCGGCTCTCTTCCATCTCGATGTCAAAGAGCCCAAGGTGGTGT 945
Db 114 GCGAGGCTGGGGGTGAGAAGGTCTCTTCCATGGAATCATCAGAGCCCTCAGTGGTGT 55
Qy 946 TTGCCAAGCAGCAGGTGGCAGACAGTGAAGTGAAGCTGAGCGGGGCCCA 997

Db 54 TTGCCAAGCAGCAGCAGCAGCAGTCAAGCAGTCCAGTGCAGGCTGTGACGGGGGCAAA 3

RESULT 4
CN302346
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

CN302346
1700532600207 GRN_ES Homo sapiens CDNA 5', mRNA sequence.
CN302346
CN302346.1 GI:47318760
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (Bases 1 to 538)
Bradenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Flisk, G.J.,
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,
Lebkowski, J. and Stanton, L.W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)
Contact: Bradenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbradenberger@geron.com
Insert Length: 538 Std Error: 0.00.
Location/Qualifiers
1..538
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and
H9"
/clone_lib="GRN_ES"
/note="Oligo dT primed, full-length enriched cDNA library
from undifferentiated hES cell lines H1 (p32), H7 (p29),
and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Query Match 26.2%; Score 262; DB 7; Length 538;
Best Local Similarity 69.5%; Pred. No. 4.5e-45;
Matches 371; Conservative 0; Mismatches 160; Indels 3; Gaps 1;

Qy 424 ATGCGTGGCTGGGGGCGGGCGGCGAGTTTGTAGTGTGAGACCTCCGAAGCCACAGCTCCACG 483
Db 1 AGGCTGAGGAGGGGCCAGTCCACATGAGCTGTGAGTGGCCCGCCAGCCAGACAGAGG 60
Qy 484 TGCACTGTGTACAGGATGGCATGGAGCTGGGCCACTCCGGTGTAGCGCTTCTTTCAGGAGG 543
Db 61 TGACGTGTGTACAGGATGGGAAGAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGGCGG 120
Qy 544 ATGTGGGACCGCGGACCGGCTGTGGCAGCCACAGTACCAGGAGGATGAGGACCT 603
Db 121 CGGGCTGATCGCGCAGCTGTGTGTGTCAGCAGGAGGCCA---GSCAGATGCTGGGGAGT 177
Qy 604 ACTCTCGCGGCTGGGGCGAGGACTCTGTGGACTTCCGGCTCCGGCTCTCTAGAGCCCAAGG 663
Db 178 ACACCTGTGAGCCCGGGGCCAGGGCTCTCTTCCACCTGGATGTTTCAAGGCCCAAGG 237
Qy 664 TGGTGTGTGTAAAGGAGCAGCTGGCAACGAGGAAGCTGCAGGACAGGACAGGAGCAGTG 723
Db 238 CGGTGTTTGAAGAGCAGCTGGCACACAGGAAGGTGCAGGCCAGGCGGGGGCATTG 297
Qy 724 CCACACTGAGCTGGAGGTGGCCAGGCCACAGACGGAGGTACGTGTTACAGGATGGA 783
Db 298 CCACTGAGCTGGAGGTGGCCAGGCCACAGACAGAGGTGACGTGGTACAGGACGGA 357
Qy 784 AGAAGCTGAGCTCCAGCTCGAAAGTGTGATGGAGGCCACAGGCTGACGCGCAGGCTGG 843

Db 358 AGAAGCTGAGCTCCAGCTCGAAAGTTTGAATGAGAGCTGTGGGTGTCACACGAGGAGCTGG 417

QY 844 TTGTGTCAGCAGGAGCCAGGCGGATGCCGGGAGTATAGCTGCAGAGGCTGGGGGCGCAGC 903

Db 418 TGGTGCAGCAGGATGCCAGGCGGACACCGGGAGTATAGCTGCGAGGCGGGGGCCAGC 477

QY 904 GGCCTCTCTTCATCTGGATGTCAAGAGCCCAAGGTGGTGTTCGCCAAGGACC 957

Db 478 GGCCTCTCTTCAGCTGGAGCTGGCAGGTCAAGTGTCTTTGTGGGCACCGAGGAGC 531

RESULT 5

AQ595824 522 bp DNA linear GSS 08-JUN-1999

LOCUS HS 2132 B2 A11 T7C CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2132 Col=22 Row=B, genomic survey sequence.

ACCESSION AQ595824

VERSION AQ595824.1 GI:5021068

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D., and Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

MEDLINE 99380589

PUBMED 10449764

COMMENT Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University Of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htsc.washington.edu
Plate: 2132 row: B column: 22
Seq primer: T7
Class: BAC ends
High quality sequence stop: 522.

FEATURES

source

1..522

Location/Qualifiers

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/clones="Plate=2132 Col=22 Row=B"

/sex="male"

/clone_lib="CIT Approved Human Genomic Sperm Library D"

/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

ORIGIN

Query Match 25.5%; Score 255.6; DB 8; Length 522;

Best Local Similarity 96.0%; Pred. No. 1e-43;

Matches 261; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 650 CTCTGAGCCCAAGTGGTGTGTTTGAAGGAGCAGCTGGCAGCAGGAAGCTGCAGGCGAGA 709

Db 251 CCCAGAGCCCAAGTGGTGTGTTTGAAGGAGCAGCTGGCAGCAGGAAGCTGCAGGAGA 310

QY 710 GGCAAGGAGCCAGTCCCACTGAGCTGCGAGTGGCCAGGCCAGACGAGGTGACGTG 769

Db 311 GGCAAGGAGCCAGTCCCACTGAGCTGCGAGTGGCCAGGCCAGACGAGGTGACGTG 370

QY 770 GTACAAGGATGGGAAGAGCTGAGCTCCAGCTCGAAAAGTGTGATGGAGGCCACAGGCTG 829

Db 371 GTACAAGGATGGGAAGAGCTGAGCTCCAGCTCGAAAAGTGTGATGGAGGCCACAGGCTG 430

QY 830 CACGCGAGGCTGCTGTGTCAGCAGGAGCCAGCGGATGCCGGGAGTATAGCTGCGA 889

Db 431 CACGCGCANGCTGTTGTGTCATCAAGCAGGCGCANGATGCCGGGAGTATATCTGCTA 490

QY 890 GGCTGGGGGGCAGCGGCTCTCTTCCCATCTGG 921

Db 491 GGCTGGGGGGCAGCGGCTCTCTTCCATCTGG 522

RESULT 6

CFI79218 621 bp mRNA linear EST 28-JUL-2003

LOCUS CFI79218

DEFINITION 813316 MARC 3PIG Sus scrofa cDNA 5', mRNA sequence.

ACCESSION CFI79218

VERSION CFI79218.1 GI:33290994

KEYWORDS EST.

SOURCE Sus scrofa (pig)

ORGANISM Sus scrofa

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

AUTHORS Smith,T.P.L., Freking,B.A., Ford,J.J., Vallet,J.L., Fox,J., Wise,T.A., Nonneman,D.J., Wray,J.E. and Keele,J.W.

TITLE A second set of porcine ESTs from a pooled-tissue normalized library

JOURNAL Unpublished (2003)

COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and trimmed with the aid of the trim_alt option. Vector identified with cross_match v0.990329.
Plate: SRG8009 row: O column: 4
Seq primer: GTAATACGACTCACTATAGG.

FEATURES

source

1..621

Location/Qualifiers

/organism="Sus scrofa"

/mol_type="mRNA"

/db_xref="taxon:9823"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 3PIG"

/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including brain, liver, muscle, placenta/endometrium, ovary, testes, and bone marrow."

ORIGIN

Query Match 24.3%; Score 243.2; DB 7; Length 621;

Best Local Similarity 66.7%; Pred. No. 4.4e-41;

Matches 363; Conservative 0; Mismatches 178; Indels 3; Gaps 1;

QY 406 TGGCGGCAGCTGTGTGGATGCGCTGGGGGCGCGGAGTTTGTAGTGTGACCT 465

Db 81 TGGCCTCTGAGTGAAGTGAAGCCATTCGCAGGCGCATTCGCAGCTGAGGTGG 140

QY 466 CCGAAGCCCAACGTCACGTGCACCTGGTACAAGGATGGCATGGAGTGGGCCACCTCCGGTG 525

Db 141 CCATGCCCATACGGAGGTGACTTGGTACAAGGATGGGAAGAGTGTGAGTTCGAGTCAA 200

QY 526 AGGGCTTCTTGAGGAGGATGTGGGACCGCGCACCGGTGGTGGCAGCCACAGTCACCA 585

Db 201 AAGTCGTGTGAGGCGCAAGGGCTGCACCCGGA---GGCTGGTGTGTCATCAAGTGGCA 257

QY 586 GGCAGGATGAAGGCACCTACTCTCTGCGGGTGGGGAGGACTCTGTGGACTTCCGGCTCC 645

Db 258 AGGCATATGCTGGGAGGTACAGCTGGAGTGTGGGGGTCTAAGGTCTCTTCCACTGG 317

QY 646 GCGTCTCTGAGCCCAAGGTGGTGTGTGTAAGGAGCAGCTGGCAGCAGGAAGCTGCAGG 705


```

Db      318 ATGTCACATATCCCTCACTGCTGTTTGCACAGGACAGCCATCAGCCACTGAAGTGCAAG 377
Qy      706 CAGAGGCGAGGAGCCAGTCCACACTGAGCTGCGAGGTGCGCCAGGCCAGAGCGGAGGTGA 765
Db      378 CCATGGCCGAGGAGGAGTCCACACTGAGCTGTGAGGTGCGCCATGCCCCAGACAGAGGTGA 437
Qy      766 CGTGTACAGGATGGGAGGAGCTGAGCTCCAGCTCGAAGTGTGATGAGGCCACAG 825
Db      438 CATGTGTAAGAGATGGGAGAGAGCTGACTTCAAGCTCAAAAGTGGGTGTGAGGCGCAATG 497
Qy      826 GCTGCACCGCAGGCTGCTGTGTGACAGCAGGCGCAGGCGGATGCCCGGGAGATAGCT 885
Db      498 GCTGCAACCGGATGCTGTGTGTGTCATCATCGGGCAATGCATATGCTGGGAGGTACAGCT 557
Qy      886 GCGAGGCTGGGGGCCAGCGGCTCTCTTCCATCTGAGTGTCAAGAGCCCAAGGTGGTGT 945
Db      558 GCGATGCTGGGGGTCATAATGCTCTCTTCCACCTGGACATCACAGAGCCCTCAGTGGTGT 617
Qy      946 TTGC 949
Db      618 TTGC 621

```

```

RESULT 7
BG680164/c
LOCUS      BG680164      1142 bp      mRNA      linear      EST 01-MAY-2001
DEFINITION 602626828F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4751671 5',
mRNA sequence.
ACCESSION  BG680164
VERSION     BG680164.1  GI:13911561
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1142)
TITLE       NIH-MGC http://mgi.nci.nih.gov/.
JOURNAL     National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT     Unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLN10608 row: h column: 08
            High quality sequence stop: 708.
            Location/Qualifiers
                1..1142
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4751671"
                /tissue_type="squamous cell carcinoma"
                /lab_host="DH10B (T1 phage-resistant)"
                /clone_lib="NCI CGAP Skn4"
                /note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
                Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
                Average insert size 1.5kb. Library constructed by Life
                Technologies. Note: this is a NCI CGAP Library."

```

FEATURES

```

source
1..1142
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4751671"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI CGAP Library."

```

ORIGIN

```

Query Match      21.7%; Score 217.6; DB 4; Length 1142;
Best Local Similarity 82.2%; Pred. No. 1.2e-35;
Matches 250; Conservative 0; Mismatches 54; Indels 0; Gaps 0;
Qy      650 CTCTCAGCCCAAGGTGGTGTTCCTAAGCAGCAGCTGCGCAGCGAGGAGTGCAGGCAGA 709
Db      314 CCCAGAGCCCAAGGCGGTGTTCCTAAGCAGCAGCAGCAGCGGAGGTGCAGTGTGA 255

```

```

Qy      710 GGCAGGAGCCAGTCCCACTGAGCTGCGAGGTGCGCCAGCCAGAGGTGACGTG 769
Db      254 AGCAGGGGCTAGGCCCACTGAGCTGTGAGGTGCGCCAGGCCAGAGAGGTGACTTG 195
Qy      770 GTACAGGATGCGGAAGAGCTGAGCTCCAGCTCGAAAGTGTGCATGGAGGCCACAGGCTG 829
Db      194 GTACAGGACGGGAAGAGCTGAGCTCCAGCTTGAAGTGTGATGGAGGCTGCAGGCTG 135
Qy      830 CACGCGCAGGCTGTTGTGACAGCGGAGGCGCAGGCGGATGCCGGGAGTATAGTTCGA 889
Db      134 TACACGGAGGCTGGTGGTGCAGCAGGCGGAGGCGCAGCACCAGAGAGTATAGCTGTGA 75
Qy      890 GGCCTGGGGGCCAGCGGCTCTCTTCCATCTCGATGTCAAGAGCCCAAGGTGGTTCG 949
Db      74 GGCCTGGGGGCCAGCAGCTCTCTTCCGCTTCAAGTGGCAGTCAAGTGGTGGGATGC 15
Qy      950 CAAG 953
Db      14 TGAG 11

```

RESULT 8

```

AW898791
LOCUS      AW898791      401 bp      mRNA      linear      EST 24-MAY-2000
DEFINITION CM0-NN0075-130400-332-f06 NN0075 Homo sapiens cDNA, mRNA sequence.
ACCESSION  AW898791
VERSION     AW898791.1  GI:8062996
KEYWORDS   EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
REFERENCE   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS     Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 401)
TITLE       Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matakuma, A., Bala, G.S., Simpson, D.H.,
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202653
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=st2-cm0-NN0075-130
400-332-f06&t3=2000-04-13&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 401.
Location/Qualifiers
1..401
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN0075"
/note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

FEATURES

```

source
1..401
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN0075"
/note="Organ: nervous normal; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```


ORIGIN

Query Match 21.7%; Score 217; DB 2; Length 401;
Best Local Similarity 85.8%; Pred. No. 1.5e-35;
Matches 241; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

QY 650 TCCTGAGCCCAAGTGGTGTGTTGCTTAAGAGCAGCTGGCAGCAGGAGCTGCAGGCAGA 709
DB 69 CCCAGAGCCCAAGGCGGTGTTCACAGGAGCAGCAGCAGGAGGTGCAGGCTGA 128
QY 710 GCGCAGAGCCAGTGCACACTGAGCTGCGAGTGGCCAGGCCAGCAGCGAGGTGACGTG 769
DB 129 GCGCGGGACCACTGCGACAGCTGAGCTGCGAGTGGCCAGGCCAGCAGCGAGGTGACGTG 188
QY 770 GTACAAGATGGAAGAACTGAGCTCCAGCTCGAAAGTGTGATGAGGAGCCACAGGCTG 829
DB 189 GTACAAGAGCGGAAGAACTGAGCTCCAGCTCGAAAGTGTGATGAGGAGCGGTGGGCTG 248
QY 830 CACGCGCAGGCTGGTGTGTCAGCAGCAGCGCAGCGGATGCGCGGGAGTATAGCTGCGA 889
DB 249 CACACGAGGCTGGTGGTGCAGGAGCGCAGCGCAGCGGAGTACAGCTGCAA 308
QY 890 GCGTGGGGCCAGCGGCTCTCTTCACCTGATGTCAAAG 930
DB 309 GCGCGGGGATCAGCGGCTGTCTTCCAGCTGCACGTGGCAG 349

RESULT 9

BQ365446
LOCUS BQ365446 260 bp mRNA linear EST 21-MAY-2002
DEFINITION RC6-GN0070-120900-022-A01 GN0070 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ365446
VERSION BQ365446.1 GI:21040958
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 260)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC6&t2=RC6-GN0070-
120900-022-A01&t3=2000-09-12&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 155
High quality sequence stop: 228.
Location/Qualifiers
1. .260
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0070"
/note="Organ: placenta_normal; Vector: puc18; Site_1:

FEATURES
source

1. .260
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0070"
/note="Organ: placenta_normal; Vector: puc18; Site_1:

ORIGIN

Query Match 21.6%; Score 216.4; DB 5; Length 260;
Best Local Similarity 95.1%; Pred. No. 1.9e-35;
Matches 234; Conservative 0; Mismatches 11; Indels 1; Gaps 1;

QY 397 ACAAGAGCATCGCGGGCAGCTGTGTGATGCCGTGGGGCCCGCCGCGAGTTTGAGT 456
DB 15 ACAAGAGCATCGCGGGCAGCTGTGCGATGACGGGCTGTAGGCCCGCGCAGTTTGAGT 74
QY 457 GTGAGACTCTCGAAGCCACGTCACGTCACCTGCTGTTCAAGGATGGAGCTGGGCC 516
DB 75 GTGAGACTCTCGAAGCCACGTCACGTCACCTGCTGTTCAAGGATGGAGCTGGGCC 134
QY 517 AC-TCCGGTCAAGCGCTTCTTTCAGGAGGATGGGACGCGGACCGGCTGGTGGCAGCC 575
DB 135 ACATCCGAGAGCGCTTCTTTCAGGAGGATGGGACGCGGACCGGCTGGTGGCAGCC 194
QY 576 ACAGTCAACGAGCAGGATGAAGCACCCTACTCTCCGCGCTGGGCGAGGACTCTGTGGAC 635
DB 195 ACAGTCAACGAGCAGGATGAAGCACCCTACTCTCCGCGCTGGGCGAGGACTCTGTGGAC 254
QY 636 TTCGGG 641
DB 255 TTCGGG 260

RESULT 10

AZ383233
LOCUS AZ383233 658 bp DNA linear GSS 02-OCT-2000
DEFINITION 1M0140L17R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0140L17 R, genomic survey sequence.
ACCESSION AZ383233
VERSION AZ383233.1 GI:10496933
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 658)
AUTHORS Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
Islan, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T.,
Reilly, M., Rose, M., Rose, R., Stokes, R., Tingey, A., von
Niederhausern, A. and Wright, D. Weiss, R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0140 row: L column: 17
Seq primer: CACACAGGAACACAGTATGACC
Class: plasmid ends
High quality sequence stop: 658.
Location/Qualifiers
1. .658
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0140L17"

FEATURES
source

1. .658
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0140L17"

COMMENT

Contact: Schageman JJ
Shohet/Garner Labs
University of Texas Southwestern Medical Center
6000 Harry Hines Blvd., N2.226, Dallas, TX 75390, USA
Tel: 214 648 1674
Email: Jeff.Schageman@UTSouthwestern.edu
cDNA library constructed by UTSW as a component of the Program for Genomic Applications (PGA) and the Reynolds Heart Disease Prevention grants for use in cDNA microarray experiments. Sequence Quality: Sequence ends were trimmed based on percentage of ambiguous base calls or 'N's in windowed segments. Sequencing: First-pass sequencing; ABI Prism 377 sequencer and analysis software.
Seq primer: M13/pUC Reverse.

FEATURES

Location/Qualifiers
1..631
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="UTSW_Smiffil"
/sex="Pooled"
/tissue_type="Diaphragm/Hind limb muscles"
/cell_type="Skeletal muscle"
/dev_stage="2 months"
/lab_host="DH5a"
/clone_lib="UTSW Adult Mouse Skeletal Muscle Library"
/note="Vector: pAMP10 (Gibco); Cloned unidirectionally. Primer: Oligo dT. RNA Isolation: cytoplasmic RNA preps (Manniat); Cloning Technique: CUA Cloning (Cloneamp, Life Technologies); Average insert size: 1.8 Kb; Insertion site: TACGTCCACTGAAATCTGAGTG----. Other information regarding entire library may be found at http://pga.swmed.edu/data/libraries/microarray_cdna_libraries.htm."

ORIGIN

Query Match 16.7%; Score 166.8; DB 4; Length 631;
Best Local Similarity 66.1%; Pred. No. 6.5e-25;
Matches 256; Conservative 0; Mismatches 128; Indels 3; Gaps 1;
Qy 168 GAGCGCAGCTGCTGGCTGAGCTGTTCAGATCAGGCTGGCGGTGTGACGTGGCTGAAGGAT 227
Db |||||
Qy 29 GCGGCACAGCTGCTGGCTGAGCTGTTCGACACAGCTGCTGTGACTTGGCTGAAGGAT 88
Db |||||
Qy 228 GGTGCACACTGTCCCGACGGCCCAAGATGAGGTGAGGATCGCGGGCGCGGGTG 287
Db |||||
Qy 89 GGTGTGAGCTGTCTCCCTAGGGCCCAAGTATGAGATGCAAGTGTGAGTGGGAGCGAGCA 148
Qy 288 CTCCTTGTGCGAGATGTGGCCGGGACGATCGAGGCTTACGAGTGGTTCAGCGCGGG 347
Db |||||
Qy 149 CTGCTGTGCGGACGTGGGACAGATGACCTGCTCTATGAGTGTGTGAGTGTGCTGGG 208
Qy 348 GGCCGATCGCTACAGCTCTCCGTGCAAGGCTCGCGGCTTCTGCAAGG---AC 404
Db |||||
Qy 209 AGCGGCACTGCTACAGCTGTGGTGTGCAAGAGCCCAAGATGTTTTCGAAGGGGCGAG 268
Qy 405 ATGCGGGGCACTGTGTGATGCGTGGTGGGGCCCGGGCGAGTTTGTAGTGTGAGACC 464
Db |||||
Qy 269 CAGGACGACGTGAGGTGAGGAGGAGCGCGGGGCCAGCGCCACACTGAGCTGGAGGTG 328
Qy 465 TCCGAAGCCCACTGCTGCACTGTGTATCAAGGATGGCATGGAGTGGGCGCACTCCGGT 524
Db |||||
Qy 329 GCCCAAGCCCACTGAGGTGACATGTTTCNAGACCGGGANNAAGCTGAGCTCCAGCTC 388
Qy 525 GAGCGCTTCTTGCAGGAGGATGTGGG 551
Db |||||
Qy 389 GAAGGTGCGCGTGGGAGCGCTCGGG 415

RESULT 13

BU439885
LOCUS 591 bp mRNA linear EST 29-NOV-2002
DEFINITION 604144795F1 CSEQRBN11 Gallus gallus cdna clone CHEST983m23 5', mRNA sequence.

ACCESSION

BU439885

VERSION

BU439885.1

KEYWORDS

EST

SOURCE

GI:25929196

ORGANISM

Gallus gallus (chicken)

REFERENCE

Gallus gallus

AUTHORS

1 (bases 1 to 591)

TITLE

Boardman,P.E., Sanz-Esquerro,J., Overton,I.M., Burt,D.W., Bosch,E.,

JOURNAL

Fong,W.T., Tickle,C., Brown,W.R.A., Wilson,S.A. and Hubbard,S.J.

MEDLINE

A Comprehensive Collection of Chicken cDNAs

COMMENT

Curr. Biol. 12 (22), 1965-1969 (2002)

PUBMED

22335534

CONTACT

Simon Hubbard

DEPARTMENT

Department of Molecular Sciences

UNIVERSITY

University of Manchester Institute of Science and Technology

(UMIST)

PO Box 88, Manchester, M60 1QD, UK

TEL

Tel: 01612008930

FAX

Fax: 01612360409

EMAIL

Email: Simon.Hubbard@umist.ac.uk.

LOCATION/QUALIFIERS

1..591

ORGANISM

/organism="Gallus gallus"

MOL TYPE

/mol_type="mRNA"

STRAIN

/strain="Layer and broiler"

DB XREF

/db_xref="taxon:9031"

CLONE

/clone="CHEST983m23"

SEX

/sex="Male and female"

TISSUE TYPE

/tissue_type="muscle"

DEV STAGE

/dev_stage="adult"

LAB HOST

/lab_host="DH103"

CLONE LIB

/clone_lib="CSEQRBN11"

NOTES

/notes="Vector: pBluescript II KS(+); Site 1: EcoRI; Site 2: NotI; This normalized library was constructed from 1 million independent clones. cDNA synthesis was initiated using an oligo(dT) primer, using methylated C in the first strand synthesis reaction. Following this first strand reaction, double-stranded cDNA was blunted, ligated to NotI adapters, digested with EcoRI, size-selected, and cloned into the NotI and EcoRI compatible sites of a custom modified MCS of the pBluescript (KS+) vector. The library was normalized in 2 rounds using conditions adapted from Soares et al., PNAS (1994) 91: 9228-9232 and Bonaldo et al., Genome Research 6 (1996): 791, except that a significantly longer reannealing hybridization was used."

QUERY MATCH

15.3%; Score 153.2; DB 5; Length 591;

BEST LOCAL SIMILARITY

64.9%; Pred. No. 4.9e-22;

MATCHES

227; Conservative 0; Mismatches 123; Indels 0; Gaps 0;

QY

652 CTGAGCCCCAAGTGTGTTTGTCTAAGGAGCAGCTGGCACGCAAGAGTTCAGGAGG 711

DB

1 CAGAGCCAGAAAGTTGTTATTTATCAACAGAGAGAGGTGCAGAGGAGGTGAAGGCTGCAC 60

QY

712 CAGAGCCAGTGCACACATGAGCTGCGAGTGGCCAGCCAGCCAGACGAGGTGACGTGGT 771

DB

61 CAACAGAAATGCCAGCTGAGCTGCGAGTGGGCGCCAGAGAGACGAGGAGTGAAGTGGT 120

QY

772 ACAAGGATGGGAGAGAGCTGAGCTCCAGCTCCAAAGTGTGCATGGAGGCCACAGGCTGCA 831

DB

121 ACAGAGAGGGGAAACTGATCACTCGAGCAAGAGTTTCAGGCTGGAGTCAGAGGGCAAC 180

QY

832 CGCGCAGGCTGTTGTGACGAGCGGACAGCCAGCCGATCCCGGGAGTATAGCTCGAGG 891

DB

181 TCGCTCGGCTGGTGTGTGAGCCAGGTGGAGAAAGATGCAGGGGAGTACACCTCGGAGG 240

QY

892 CTGGGGCCAGCGGCTCTCTTCCATCTCGATGTCAAGAGCCCAAGGCTGTTGTTGCCA 951

Db 241 CTGTGGCCAGAACTGACCTTCAAGATTGATGTACAGAGCCAAACCTGTGTTTATAA 300
 QY 952 AGGACCAGGTGGCCACACAGTAGTGAGGTGCGAGGCTGAGGCGGGGCCCAATGCC 1001
 Db 301 ACCAGGAAAGGTGCAGAGGGAGGTGAGTGTCTCTGCGCAAAAGTGCC 350

RESULT 14
 BE159678/c
 LOCUS BE159678 343 bp mRNA linear EST 21-JUN-2000
 DEFINITION MRO-HT0407-180400-015-e07 HT0407 Homo sapiens cDNA, mRNA sequence.
 ACCESSION BE159678
 VERSION BE159678.1 GI:8622399
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 343)
 AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed

JOURNAL sequence tags
 MEDLINE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 PUBMED 20202663
 COMMENT 10737800

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001

Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml.pl?l1=6t2-MRO-HT0407-180
 400-015-e07&t3=2000-04-18&t4=1)

Seq primer: puc 18 forward

High quality sequence stop: 343.

Location/Qualifiers

1..343
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT0407"

/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 14.9%; Score 149.4; DB 2; Length 343;
 Best Local Similarity 88.7%; Pred. No. 3e-21;
 Matches 173; Conservative 0; Mismatches 21; Indels 1; Gaps 1;

QY 650 CTCTAGCCCAAGGTGGTGTGGCTTAAGGACGAGCTGGCACGACGAAAGCTGCAGGCAGA 709
 Db 228 CCCAGAGCCCAAGGTGGTGTGGCTTAAGGACGAGCTGGCACGACGAAAGCTGCAGGCAGA 169

QY 710 GCGAGGACCCAGTGCACACTGAGCTGGAGTGGCCAGGCCACAGACGAGGAGTGCACGTG 769
 Db 168 GCGCGGGGCCAGTGCACCGCTGAGCTGCAGAGTGGCCAGGCCACAGACGAGGAGTGCACGTG 109

QY 770 GTACAGGATGGGAAGAAGCTGAGCTCCAGCTCGAAAGTGTGCATGGAGGCCACAGGCTG 829

Db 108 GTACAGGATGGGAAGAAGCTGAGTTCAGCTCGAAAGTGGCGGTGGA-GCGGTGGCTG 50
 QY 830 CACGCGCAGGCTGGT 844
 Db 49 CACACGGAGGCTGGT 35

RESULT 15
 BZ837763
 LOCUS BZ837763 397 bp DNA linear GSS 18-MAR-2003
 DEFINITION CH240_248L16_TV CHORI-240 Bos taurus genomic clone CH240_248L16,
 genomic survey sequence.

ACCESSION BZ837763
 VERSION BZ837763.1 GI:29065122
 KEYWORDS GSS.

SOURCE Bos taurus (cow)

ORGANISM Bos taurus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.

REFERENCE 1 (bases 1 to 397)
 AUTHORS Zhao,S., Shetty,J., Shatsman,S., Tsengaye,G., Geer,K.,
 Shvartsbeyn,A., Gebregorgis,E., Chen,D., Riggs,P., de Jong,P.,
 Crawford,A.M. and McEwan,J.C.
 Bovine BAC End Sequences from Library CHORI-240
 Unpublished (2003)

TITLE JOURNAL

COMMENT

Contact: Shaying Zhao
 Department of Eukaryotic Genomics
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: szhao@tigr.org

Clones are derived from the bovine BAC library CHORI-240
 (http://www.chori.org/bacpac/bovine240.htm). For BAC library
 availability, please contact Pieter de Jong (pjejong@mail.cho.org).
 Clones may be purchased from BACPAC Resources
 (http://www.chori.org/bacpac/ordering_information.html). This work
 was undertaken as part of the International Bovine BAC Mapping
 Consortium (IBBMC) by AgResearch Ltd., New Zealand and The
 Institute of Genomic Research (TIGR), USA.

Plate: 248 row: L column: 16

Seq primer: T7

Class: BAC ends.

Location/Qualifiers

1..397
 /organism="Bos taurus"
 /mol_type="genomic DNA"
 /strain="bred: Hereford"
 /db_xref="taxon:9913"
 /clones="CH240_248L16"
 /sex="Male"
 /cell_type="Blood"
 /clone_lib="CHORI-240"

/note="vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;
 Hereford bull L1 Domino 99375; CHORI-240 Bovine BAC
 library (Male) produced by Pieter de Jong"

ORIGIN

Query Match 14.6%; Score 146.6; DB 8; Length 397;
 Best Local Similarity 85.0%; Pred. No. 1.2e-20;
 Matches 164; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 650 CTCTAGCCCAAGGTGGTGTGGCTTAAGGACGAGCTGGCACGACGAAAGCTGCAGGCAGA 709
 Db 205 CCCAGAGCCCTCGGTGGTGTGGCTTAAGGACGAGCTGGCACGACGAAAGCTGCAGGCCTG 264

QY 710 GCGAGGACCCAGTGCACACTGAGCTGGAGTGGCCAGGCCACAGACGAGGAGTGCACGTG 769
 Db 265 GCGCGGACCAAGTGCACCGCTGAGCTGCAGAGTGGCCAGGCCACAGACGAGGAGTGCATGTG 324

QY 770 GTACAGGATGGGAAGAAGCTGAGCTCCAGCTCGAAAGTGTGCATGGAGGCCACAGGCTG 829

Db	325	GTACAAAGGACGGCAAGAGCTGAGTTTCGAGCTCGAAAGTTCGTGTGAGGCCACGGGCGG	384
QY	830	CACGGCGAGGCTG	842
Db	385	CGGGCGGCGGCTG	397

Search completed: March 21, 2005, 14:13:54
Job time : 3665.47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 16:50:48 ; Search time 10983.1 Seconds
(without alignments)
10400.615 Million cell updates/sec

Title: US-10-077-130-6_COPY_19500_22500
Perfect score: 3001
Sequence: 1 cctacggagcagaactcggg.....ccacctggagcaaacagcga 3001

Scoring table: IDENTITY_NUC
Gap 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST.*

1: gb_est1.*

2: gb_est2.*

3: gb_hic.*

4: gb_est3.*

5: gb_est4.*

6: gb_est5.*

7: gb_est6.*

8: gb_g881.*

9: gb_g882.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	867	28.9	907	5	BX460671
2	541	18.0	622	4	BG999843
3	482.2	16.1	723	7	CK357545
4	434	14.5	517	4	BG989614
5	307	10.2	920	2	BF160452
6	281.6	9.4	429	1	AI604658
7	279	9.3	924	5	BX409792
8	267.2	8.9	311	2	BF855491
9	259	8.6	570	4	BI345921
10	254.4	8.5	929	5	BX409793
11	246	8.2	397	4	BG986727
12	229.6	7.7	371	1	AJ573174
13	229.4	7.6	368	2	AW425352
14	208.2	6.9	653	6	BY739930
15	197.2	6.6	465	2	BE757677
16	190.8	6.4	771	7	CK601206
17	190.6	6.4	447	2	BB848582
18	190	6.3	811	7	CK599424
19	185.8	6.2	389	5	BY088395
20	185	6.2	236	2	BF856372
21	180.6	6.0	297	1	AA771462
22	172	5.7	976	3	BC012210
23	171.6	5.7	580	4	BI343159
24	170.4	5.7	797	4	BI663166

25	167.8	5.5	708	2	AW476323
26	163.8	5.5	972	4	BI554559
27	163.6	5.5	982	4	BF976537
28	160.8	5.4	663	9	AG112651
29	159.2	5.3	237	2	BB605899
30	158.8	5.3	663	2	BF019568
31	158.6	5.3	671	2	BF727181
32	158.6	5.3	869	7	CR553384
33	156.8	5.2	657	4	BM763093
34	156.6	5.2	844	7	CF617929
35	155.8	5.2	636	1	AA557328
36	155.2	5.2	526	5	BQ554403
37	155	5.2	1058	4	BG421064
38	154.4	5.1	997	5	BQ071271
39	153.4	5.1	887	7	CN646845
40	152.8	5.1	614	6	CB372576
41	152.2	5.1	1039	4	BG419640
42	152	5.1	718	7	CN789380
43	150.2	5.0	576	5	BQ187532
44	150.2	5.0	978	7	CN646611
45	149	5.0	556	7	CO503963

ALIGNMENTS

RESULT 1
BX460671
LOCUS
DEFINITION BX460671 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CS0DF017YK06 5-PRIME, mRNA sequence.
ACCESSION BX460671
VERSION BX460671.2 GI:47053688
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 907)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 22, 2003 this sequence version replaced gi:31023260.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with NotI and cloned
into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1025.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0DF017BF03QPI&c=1025.r.
Location/Qualifiers
1. 907
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF017YK06"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"

FEATURES

source
1. 907
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF017YK06"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN

Query Match 28.9%; Score 867; DB 5; Length 907;
Best Local Similarity 99.2%; Pred. No. 6.1e-164;

Matches 900; Conservative 2; Mismatches 2; Indels 3; Gaps 3;									
Qy	1903	CTCTCCAGGTGGGACAGAGCCTGGCCCTCCCTGGATGGAGGGGTGGACCCAGGA	1962						
Db	1	CTCTCCAGGTGGGACAGAGCCTGGCCCTCCCTGGATGGAGGGGTGGACCCAGGA	60						
Qy	1963	GGCTGAGGATCTCCGATCCACACCCACTTTCAGCGGCGCTCAGGAACAGGCGACCAT	2022						
Db	61	GGCTGAGGATCTTCGCACTCCACACCCACTTTCAGCGGCGCTCAGGAACAGGTCACAT	120						
Qy	2023	CGCAGAGTCTCCCTGGTGGTTCGCGGGGCTACGCGAGGCGTGGCTGATGCGACCTT	2082						
Db	121	CGCAGAGTCTCCCTGGTGGTTCGCGGGGCTACGCGAGGCGTGGCTGATGCGACCTT	180						
Qy	2083	TGCTTTTGGTGGAGATCAGGGGGCATGCTGGGGCAGGGGCCCATGTGGGCCAGGATAGC	2142						
Db	181	TGCTTTTGGTGGAGATCAGGGGGCATGCTGGGGCAGGGGCCCATGTGGGCCAGGATAGC	240						
Qy	2143	CTGGGCTGTCTCCAGTCGAGGAGGAGGAGCAGGAGGCGCCAGGGCTGAGTCCCGATC	2202						
Db	241	CTGGGCTGTCTCCAGTCGAGGAGGAGGAGCAGGAGGCGCCAGGGCTGAGTCCCGATC	300						
Qy	2203	GGAGGACAGCAGGAGGCCAGGGCTGAGAGCCACTGCCCCAGGTCAGTCAAGGCGCTGT	2262						
Db	301	GGAGGACAGCAGGAGGCCAGGGCTGAGAGCCACTGCCCCAGGTCAGTCAAGGCGCTGT	360						
Qy	2263	GCCTGAGGTGGCAGGCTCCACAGAGCTCTCCAGAGGCCACCCCATGGGAGGACAT	2322						
Db	361	GCCTGAGGTGGCAGGCTCCACAGAGCTCTCCAGAGGCCACCCCATGGGAGGACAT	420						
Qy	2323	CGGCGAGTCTCCCTGGTGCAGATCCGGGACCTGTCAAGTGATGCGGAGGCGGCCGAC	2382						
Db	421	CGGCGAGTCTCCCTGGTGCAGATCCGGGACCTGTCAAGTGATGCGGAGGCGGCCGAC	480						
Qy	2383	AATATCCCTGGAATTTCCGAGGTGGACCCCGCCTACTCACTCTCAGACCTGTACGA	2442						
Db	481	AATATCCCTGGAATTTCCGAGGTGGACCCCGCCTACTCACTCTCAGACCTGTACGA	540						
Qy	2443	TATCAAGTACCTCCATTCGAGTTTATGATCTTCAGGAAGTCCCAAGTCCGCTCAGCC	2502						
Db	541	TATCAAGTACCTCCATTCGAGTTTATGATCTTCAGGAAGTCCCAAGTCCGCTCAGCC	600						
Qy	2503	AGAGCGGCTCTCCCATGCTGAGGAGGAGTGGCCGAGTTCGCGAGGCCAGCTGGCC	2562						
Db	601	AGAGCGGCTCTCCCATGCTGAGGAGGAGTGGCCGAGTTCGCGAGGCCAGCTGGCC	660						
Qy	2563	CTGGCCAGGTGAATCTGGGCCCCCAGCAGGCGCT-GGAGATCAGAGAGGTACAGGATG	2621						
Db	661	CTGGCCAGGTGAATCTGGGCCCCCAGCAGGCGCT-GGAGATCAGAGAGGTACAGGATG	720						
Qy	2622	TGGACGGCTCTGTCGAGAGGCTGCGGTGGGAGGAGCGAAGTGGTCTCTCGCGCTCAC	2681						
Db	721	TGGACGGCTCTGTCGAGAGGCTGCGGTGGGAGGAGCGAAGTGGTCTCTCGCGCTCAC	780						
Qy	2682	GCAGCCTCTTCCACTTCCC-TGGGAGGCACTGGCGCTGGATGAGCTTCGAGAGCT-GGG	2739						
Db	781	GCAGCCTCTTCCACTTCCC-TGGGAGGCACTGGCGCTGGATGAGCTTCGAGAGCTGGGG	840						
Qy	2740	GCTGGCTGAGAGTGAAGGCTCCGTGGAGCATCTCCCGGATCTCTGAAGGCGAGGCC	2799						
Db	841	GCTGGCTGAGAGTGAAGGCTCCGTGGAGCATCTCTCCCGATCTCTGAAGGCGAGGCC	900						
Qy	2800	GGAGGT 2806							
Db	901	GGAGGT 907							
RESULT 2									
BG999843/c									
LOCUS									
DEFINITION									
MR3-HN0063-030101-001-h03 HN0063 Homo sapiens cDNA, mRNA sequence.									
ACCESSION									
BG999843									
VERSION									
BG999843.1 GI:14403915									

KEYWORDS	EST.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
	1 (bases 1 to 622)
	Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
	Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
	Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
	Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
	O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
	Simpson,A.J.
TITLE	Shotgun sequencing of the human transcriptome with ORF expressed
	sequence tags
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE	20202663
PUBMED	10737800
COMMENT	Contact: Simpson A.J.G.
	Laboratory of Cancer Genetics
	Ludwig Institute for Cancer Research
	Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
	Brazil
	Tel: +55-11-2704922
	Fax: +55-11-2707001
	Email: asimpson@ludwig.org.br
	This sequence was derived from the FAPESP/LICR Human Cancer Genome
	Project. This entry can be seen in the following URL
	(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR3&t2=MR3-HN0063-
	030101-001-h03&t3=2001-01-03&t4=1)
	Seq primer: puc 18 forward
	High quality sequence stop: 621.
FEATURES	Location/Qualifiers
source	1..622
	/organism="Homo sapiens"
	/mol_type="mRNA"
	/db_xref="taxon:9606"
	/dev_stage="Adult"
	/clone_lib="HN0063"
	/note="Organ: head normal; Vector: puc18; Site 1: SmaI;
	Site 2: SmaI; A mini-library was made by cloning products
	derived from ORESTES PCR (U.S. Letters Patent application
	No. 196,716 - Ludwig Institute for Cancer Research)
	profiles into the pUC 18 vector. Reverse transcription of
	tissue mRNA and cDNA amplification were performed under
	low stringency conditions."
ORIGIN	
	Query Match 18.0%; Score 541; DB 4; Length 622;
	Best Local Similarity 98.8%; Pred. No. 2.6e-98;
	Matches 566; Conservative 0; Mismatches 5; Indels 2; Gaps 2;
Qy	2235 CACTGCCCAAGGTCAAGGCTGTGCTGAGTGGCAGGCTCCACCA-CGAGC 2293
Db	622 CACTGCCCAAGGTCA-TCCAAAGGCTGTCTGAGGTCCGACAGGCTCCACCAAGGAGC 564
Qy	2294 TCTCCAGAGCCACCCCATGGAGGACATCGGCAGGCTCTCCCTGTGTCAGATCCGGAC 2353
Db	563 TCTCCAGAGCCACCCCATGGAGGACATCGGCAGGCTCTCCCTGTGTCAGATCCGGAC 504
Qy	2354 CTGTCAAGTGTGTCGAGGCGCGGACACATAATCTCCGAGATTTCCGAGGTGACCCC 2413
Db	503 CTGTCAAGTGTGTCGAGGCGCGGACACATAATCTCCGAGATTTCCGAGGTGACCCC 444
Qy	2414 GCCTACCTCAACCTCTCAGACCTGTAGCATATCAGGTACCTCCATTCGAGTTTATGATC 2473
Db	443 GCCTACCTCAACCTCTCAGACCTGTAGCATATCAGGTACCTCCATTCGAGTTTATGATC 384
Qy	2474 TTCAGGAAGTCCCAAGTCCGCTCAGCCAGAGCCGCTCCGCCCATGGCTGAGGAGGAG 2533
Db	383 TTCAGGAAGTCCCAAGTCCGCTCAGCCAGAGCCGCTCCGCCCATGGCTGAGGAGGAG 324
Qy	2534 CTGGCCGAGTTCGCGAGGCCACAGTGGCCCTGGCCAGGTGAATCTGGGCCCCCAGCGAGC 2593

Db 323 CTGGCCGAGTTCCTCCGAGCCCACTGGCCCTGGCCAGGTGAACCTGGGCCCCCAGCCAGGC 264
Qy 2594 CTGAGATATCAGAGGAGTCAAGAGATGTGACGCGCTGCTGGAGAGGCTGCGGTGGGC 2653
Db 263 CTGAGATATCAGAGGAGTCAAGAGATGTGACGCGCTGCTGGAGAGGCTGCGGTGGGC 204
Qy 2654 AGGAAGCGCAAGTGTCTCCCGCTCAGCAGAGCTCTTCCACTTCCCTGGGAGGACCTG 2713
Db 203 AGGAAGCGCAAGTGTCTCCCGCTCAGCAGAGCTCTTCCACTTCCCTGGGAGGACCTG 144
Qy 2714 CCGCTGGATGAGCCTCGAGAGCTGGGCTGGCTGGAGAGTGAAGGCTCCGCTGGAGCAC 2773
Db 143 CCGCTGGATGAGCCTCGAGAGCTGGGCTGGCTGGAGAGTGAAGGCTCCGCTGGAGCAC 84
Qy 2774 ATCTCCCGGATCTCGAAGGAGGCGCGGAAGT 2806
Db 83 ATCTCCCGGATCTCGAAGGAGGCGCGGAAGT 51

RESULT 3

CK357545 723 bp mRNA linear EST 23-DEC-2003.
LOCUS AGNCOURT 17184642 NIH MGC 233 Rattus norvegicus cDNA clone
IMAGE:7105047 5', mRNA sequence.

CK357545
CK357545.1 GI:40323477

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Rattus norvegicus (Norway rat)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Rattus.

1 (bases 1 to 723)

NIH-MGC <http://mgi.nci.nih.gov/>.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgapbs-re@mail.nih.gov

Tissue Procurement: Howard Jacobs

cDNA Library Preparation: Express Genomics

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

<http://image.llnl.gov>

Plate: LLNL4963 row: 1 column: 13

High quality sequence stop: 724.

Location/Qualifiers

1..723

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clones="IMAGE:7105047"

/tissue type="heart, pooled"

/lab host="DH10B Tona"

/clone lib="NIH MGC 233"

/note="Organ: heart; Vector: pExpress-1; Site 1: EcoRV;

Site 2: NotI; RNA obtained from pooled heart tissue from a

mix of male and female animals at 8 wk old. Tissues were

snap-frozen and kept at -80C for two days before RNA

extraction and purification (Tri-reagent method). cDNA was

primed using oligo-dT primer:

5'-pCACTAGTCTTATGTCGGAGCGGGCGGCTT-3' and cloned into

the EcoRV/NotI sites of pExpress-1. Size-selection

resulted in an average insert size of 2 kb. This primary

library is not normalized (normalized primary library is

NIH MGC 234) and was constructed by Express Genomics

(Frederick, MD). Note: this is a NIH_MGC library."

(Frederick, MD). Note: this is a NIH_MGC library."

ORIGIN

COMMENT

Contact: Simpson A.J.G.

Contact: Simpson A.J.G.

Contact: Simpson A.J.G.

Contact: Simpson A.J.G.

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=MR2&t2=MR2-Hrl160-
 190101-013-ell&t3=2001-01-19&t4=1)
 Seq primer: puc 18 forward
 High quality sequence stop: 355.

FEATURES

Location/Qualifiers
 1..517
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT1160"

/note="Organ: head neck; Vector: puc18; Site_1: SmaI;
 Site_2: SmaI; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 14.5%; Score 434; DB 4; Length 517;
 Best Local Similarity 94.5%; Pred. No. 8.8e-77;
 Matches 482; Conservative 0; Mismatches 25; Indels 3; Gaps 3;

Qy 266 AAGCCCTCAACATCTGTGTCATCTCTCCCGGGAAGACATTAAATCTGGACTTT 325
 |||||
 Db 508 AAGCCCTCAACATCTGTGTCATCTCTCCCGGGAAGACATTAAATCTGGACTTT 450
 |||||
 Qy 326 GGCTTTGCCAGAACATCACCCAGCAGAGCTGCAGTTTCAGCCAGTACGGCTCCCTCTGAG 385
 |||||
 Db 449 GGCTTTGCCAGAACATCACCCAGCAGAGCTGCAGTTTCAGCCAGTACGGCTCCCTCTGAG 390
 |||||
 Qy 386 TTGCTGT-CCCCGAGATCATTCAGCAGAACCCCTGTGAGCGAAGCCCTCCGACATTTGGGC 444
 |||||
 Db 389 TTGCTGTCCCCCGAGATCATTCAGCAGAACACCTGTGAGCGAAGCCCTCCGACATTTGGGC 330
 |||||
 Qy 445 CATGGGTGTATCTCTACCTCAGCTGACCTGCTCATCCCCATTTTCGCGCGAGAGTGA 504
 |||||
 Db 329 CATGGGTGTATCTCTACCTCAGCTGACCTGCTCATCCCCATTTTCGCGCGAGAGTGA 270
 |||||
 Qy 505 CCGTGCCAACCTCTCTGAACCGTCTGGAGGGGCGGTGTCTATGGAGCAGCCCATGGCTGC 564
 |||||
 Db 269 CCGTGCCAACCTCTCTGACGCTCTGGAGGGGCGGTGTCTATGGAGCAGCCCATGGCTGC 210
 |||||
 Qy 565 CCACCTCAGCGAAGACGCCAAGACTTCATCAAGGCTACGCTGCAGAGAGCCCTCAGGC 624
 |||||
 Db 209 CCACCTCAGCGAAGACGCCAAGACTTCATCAAGGCTACGCTGCAGAGAGCCCTCAGGC 150
 |||||
 Qy 625 CCGGCTAGTCCGCGCCAGTGCCTCTCCACCCCTGCTTCTGAAATCCATGCTCGCGGA 684
 |||||
 Db 149 CCGGCTAGTGTGGCCAGTGCCTCTCCACCCCTGCTTCTGAAATCCATGCTCGCGGA 90
 |||||
 Qy 685 GGAGGCCCATTTTCATCAACACCAAGCAGCTCAAGTTCTCTGTGCGCCGAAAGTCCCTGGCA 744
 |||||
 Db 89 GGAGGCCCATTTTCATCAACACCAAGCAGCTCAAGTTCTCTGTGCGCCGAA-TCACTGGCA 31
 |||||
 Qy 745 GCGTTCCTGATGAGCTACAGTCCATCTT 774
 |||||
 Db 30 GCGTTGCTGAATGAGCCCAAGGATCCT 1
 |||||

RESULT 5
 BF160452
 LOCUS

920 bp mRNA linear EST 30-OCT-2000

DEFINITION 601768424F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3988060 5',
 mRNA sequence.

ACCESSION BF160452
 VERSION BF160452.1 GI:11040468
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 920)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.

CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINL at:
 http://image.llnl.gov
 Plate: LL49195 row: 0 column: 05
 High quality sequence stop: 723.

FEATURES

Location/Qualifiers
 1..920

/organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C2BECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:3988060"
 /tissue="spontaneous tumor, metastatic to mammary."
 Stem cell origin.
 /lab_host="DH10B"
 /clone_lib="NCI_CGAP_Lu29"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH"

ORIGIN

Query Match 10.2%; Score 307; DB 2; Length 920;
 Best Local Similarity 74.4%; Pred. No. 3.5e-51;
 Matches 474; Conservative 0; Mismatches 140; Indels 23; Gaps 6;

Qy 2277 GGCTCCACACGAGCTCTCCAGAGCCACCCCATGGGAGGACATCGGGCAGGTCTCCC 2336
 |||||
 Db 3 GGTTTCCCTAAGACCTCTCCAGGCTCACCCCTTGGGAGGAAGTTGACAGGTTTCCC 62
 |||||
 Qy 2337 TGGTCAGATCCGGGACCTGTCTAGGTGATGCGAGGCGCGCACAAATATCCCTGGACA 2396
 |||||
 Db 63 TGGTCAGATCCGGGATCTGTCTGGTGATGCGGAAGCAGCTGACATATCTCTTGGACA 122
 |||||
 Qy 2397 TTTCCGAGGTGACCCGGCTACTCAACCTCTCAGACCTGTACGATATCAAGTACCTCC 2456
 |||||
 Db 123 TTTCTGAGGTGATCTCTGCTTCAACCTCTCGGATCTATATGACATCAAAATATCTCC 182
 |||||
 Qy 2457 CATTCGAGTTTATGATCTTCAGGAAGTCCCCAAAGTC---CGCTCAGCAGAGCGCCCT 2513
 |||||
 Db 183 CATTTGAGTTCAATGATCTTCAGAAAGGTGCCAAACCTATAGAGCAGCAGAGTCACTG 242
 |||||
 Qy 2514 CC-----CCCATGGCTGAGGAGAGCTGGCCGAGTTTCCCGAG--CCACAGTGGCCCTG 2565
 |||||
 Db 243 GCTCAGAAACCGAAGCTGGGCAAGGCTGGCAGATTTCCTGGAGGAGCGCATGGCCCT 302
 |||||
 Qy 2566 GCCAGGTGAATCTGGGCCCCCACCAGCGCCTGGGAGATCACAGGAGTGCAG-----G 2617
 |||||
 Db 303 GGCAGGCGAGCTGGGACTTCGTCTGTCTGAGATTACAGAGGAGCGCGAGAGCCAGG 362
 |||||
 Qy 2618 GATGTGACGCGCTGTCTGGCAGAGGCTGCGTGGGAGGAGCGCAAGTGGTCTCTCGCCG 2677
 |||||
 Db 363 GACCTGGAAGCACTGTCTGGCGAGGCTGCTGTGGCAGGAAGCGCAAGTGGTCC---CCC 419
 |||||


```

/lab host="DH10B"
/clone_lib="MARC 2PIC"
/note="vector: pCMV SPORT6; Site1: NotI; Site2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

ORIGIN
Query Match      8.6%; Score 259; DB 4; Length 570;
Best Local Similarity 69.5%; Pred. No. 1.5e-41;
Matches 405; Conservative 0; Mismatches 130; Indels 48; Gaps 2;

Qy 1900 GGGCTCTCCAGGTGGGCACAGAGCCCTGGCCCTCCCTGGATGGGAGGCTGGACCCA 1959
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 27 GGCAGCCCCCAATAGCAGGAGCCAGGCCCTCCCTGGATGCTGAGGCGCTTACCCA 86
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1960 GGAGGCTGAGGATCTGTCGACTCCACACCACCTTGACGGGGCTCAGGAACAGGGAC 2019
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 87 GGAGACGGAACCTGTCTAGCTTTCACCTGGTCCACAGCGGCTCAGGAACAGGGAC 146
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2020 CATGGCAAGTTCTCCCTGGGTGTCGGGGGGCTACGACGGCTGGCTGCTATGGAC 2079
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 147 CACACGGAAGTTCTCTTGGGATACCGTGGGGGCTATGACAGGGTGGCGGCTACGGCAC 206
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2080 CTTTGCCCTTTCGTGAGATGACGGGGCATGTCGGGGCAGGGCCATGTCGGCCAGGAT 2139
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 207 CTTTGCCCTTTCGTGGGATGCTGGGGCATGCTGGGGCAGGACCTGTGGGCCAAGAT 266
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2140 AGCTGGGCTGTGTCAGTCCGAGGAGGAGGAGCAGGAGGAGCCAGGGCTGAGTCCCA 2199
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 267 GGCCTGGGCCCTCTCAGTC----- 287
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2200 GTCGAGAGAGCAGAGAGGCGGCTGAGAGCCCACTGCCCCAGGTGCTGCAAGGCC 2259
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 288 CTCAGAGAGCAGAGAGGACCAAGGGGCCAGAGCCACACCCAGGCGCACACAGTCCC 347
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2360 TGTGCTGAGGTGCGCAGGGTCCACACAGAGGCTCTCCAGAGCCACCCCATGGAGGA 2319
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 348 CTAAGTGGAGGTGGCAAGGCCCTCTGAGGGCTCCAGAGATGAGCTCATGGAGGA 407
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2320 CAT-----CGGGCAGGTCTCCCTGGTGCAGATCCGGACCTGTGAGGTGATGCGGA 2370
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 408 CTTAAGTGGAGGCTCCAGAGGTGCTGCTGTACAGATCGGGACCTGTCTCATGGACCGGA 467
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2371 GCGCGCCGACACATATCCCTGGACATTCGAGGTGAGACCCGCTACCTCAACCTCTC 2430
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 468 AGCAGCTGACACTGTATCCCTGGATATCTCAGAGGTGGATCTGCTCACTCAACCTGTC 527
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2431 AGACTGTACGATATCAAGTACCTCCCATTCGAGTTTATGATC 2473
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 528 TGACTGTATGACATCAATACCTGCCATTTGAATTATGATC 570
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 10
LOCUS      BX409793
DEFINITION BX409793 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
ACCESSION  BX409793
VERSION     BX409793.2 GI:46926497
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 527)
AUTHORS    Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL    Unpublished (2001)
COMMENT    On May 13, 2003 this sequence version replaced gi:30640364.
            Contact: Genoscope
            Genoscope - Centre National de Sequencage
            2 rue Gaston Cremieux, CP 5705 - 91057 EVRY cedex - FRANCE
            Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

```

1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.

This sequence belongs to sequence cluster 1025.r

For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS05BAP014ZC08_AF01308_2&c=1025.r

FEATURES

Location/Qualifiers
1..929
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF017YK06"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

ORIGIN

```

Query Match      8.5%; Score 254.4; DB 5; Length 929;
Best Local Similarity 94.1%; Pred. No. 1.4e-40;
Matches 273; Conservative 0; Mismatches 15; Indels 2; Gaps 1;

Qy 2712 TGCCGCTGGATGAGCCCTGCAGAGCTGGGGCTGCGTGAGAGAGTGAAGSCCTCCGTGGAGC 2771
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1 TGCCTGTCGCGAACCTGTCANGACTGGGGCTGCGTGNGNGTGAAGSCCTCCGTGGAGC 60
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2772 ACATCTCCCGATCTTGAAGGGCAGCGGCGAAGGTCTGGAGAAGAGGGGCCCCCAGGA 2831
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 61 ACATCTCCCGATCTTGA--GGGCAGGCGCGAGGTCTGGAGAAGAGGGGCCCCCAGGA 118
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2832 AGAAGCAGGCTTGTCTCTTCCGGCTCTCAGGTCTGAAGAGCTGGACCGAGGCGCGA 2891
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 119 AGAAGCAGGCTTGTCTTCCGGCTCTCAGGTCTGAAGAGCTGGACCGAGGCGCGA 178
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2892 CATTCTTAAGGAGCTCTCAGATGAGACTGTGGTCTTGGGCGAGTCACTGACCTGGGCT 2951
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 179 CATTCTTAAGGAGCTCTCAGATGAGACTGTGGTCTTGGGCGAGTCACTGACCTGGGCT 238
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 2952 GCCAGGTGTGAGCCCGAGCCAGCTCCCGAGGCGACCTCGAGCAAGAGACGGA 3001
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 239 GCCAGGTGTGAGCCCGAGCCAGCTCCCGAGGCGACCTCGAGCAAGAGACGGA 288
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 11

```

LOCUS      BG986727
DEFINITION BG986727 PM1-HT1170-030101-001-g12 HT1170 Homo sapiens cDNA, mRNA sequence.
ACCESSION  BG986727
VERSION     BG986727.1 GI:14390797
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
            1 (bases 1 to 397)
AUTHORS    Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
            Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,
            Goldman, G.H., Carvalho, A.F., Mateukuma, A., Baia, G.S., Simpson, D.H.,
            Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
            O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
            Simpson, A.J.
            Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
            Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
            20202663
            MEDLINE
            10737800

```


TITLE	Oligo-directed RNase H cleavage of abundant mRNAs in skeletal muscle					
JOURNAL COMMENT	Unpublished (2003) Contact: Laveder P CRIBI Biotechnology Centre University of Padua Via U. Bassi 58/B, 35121 Padua, Italy ABI Chromatograms and other information are available on WWW at http://muscle.cribi.unipd.it BIOLIMS code: shr-000006-0-A07 Seq primer: PC2R. Location/Qualifiers 1..371 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="HSPD46382" /sex="female" /tissue_type="pectoral muscle (after mastectomy)" /clone_lib="HMJ/RH2"					
FEATURES source						
ORIGIN						
Query Match	7.7%; Score 229.6; DB 1; Length 371;					
Best Local Similarity	94.4%; Pred. No. 1.2e-35;					
Matches 238; Conservative	0; Mismatches 14; Indels 0; Gaps 0;					
QY	2277	GGGCTCCACACGAGGAGCTCTCCAGAGGCCACCACCCATGGGAGGACATCGGCAGGTCTCCC	2336			
Db	9	GAGATGGCACAGAGTGAATGCAGAGGCCCAACCCATGGGAGGACATCGGCAGGTCTCCC	68			
QY	2337	TGGTGCAAGATCCGGGACCTGTTCAGGTGATCGGAGCGCGCCGACACAATATCCTCGACA	2396			
Db	69	TGGTGCAAGATCCGGGACCTGTTCAGGTGATCGGAGCGCGCCGACACAATATCCTCGACA	128			
QY	2397	TTTCCGAGGTGACCCCGGCTACTCAACTCTCAGACCTGTACGATATCAAGTAGCTCC	2456			
Db	129	TTTCCGAGGTGACCCCGGCTACTCAACTCTCAGACCTGTACGATATCAAGTAGCTCC	188			
QY	2457	CATTTCGAGTTTATGATCTTCAGAAAGTCCCCAAGTCCGCTCAGCCAGAGCCGCCCTCCC	2516			
Db	189	CATTTCGAGTTTATGATCTTCAGAAAGTCCCCAAGTCCGCTCAGCCAGAGCCGCCCTCCC	248			
QY	2517	CCATGGCTGAGG	2528			
Db	249	CCATGGCTGAGG	260			
RESULT 13						
AW425352	368 bp mRNA linear EST 25-APR-2001					
LOCUS	57199 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.					
DEFINITION	AW425352					
ACCESSION	AW425352.1 GI:69532399					
VERSION	EST.					
KEYWORDS	Bos taurus (cow)					
SOURCE	Bos taurus					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.					
REFERENCE AUTHORS	Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Pertea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keele,J.W.					
TITLE	Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle					
JOURNAL MEDLINE PUBMED	Genome Res. 11 (4), 626-630 (2001) 21180013 11282978					
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA					

Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
 Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome
 Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l1=PM1&l2=PM1-HT1170-
 030101-001-g12&t3=2001-01-03&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 4
 High quality sequence stop: 320.
 Location/Qualifiers
 1..397
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HT1170"
 /note="Organ: head neck; Vector: puc18; Site: 1: SmaI;
 Site 2: SmaI; A mini-library was made by cloning products
 derived from ORESPEC PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the puc 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN
 Query Match 8.2%; Score 246; DB 4; Length 397;
 Best Local Similarity 92.4%; Pred. No. 6.1e-39;
 Matches 281; Conservative 0; Mismatches 20; Indels 3; Gaps 2;
 QY 30 ACAGGGAGCGAGACATCTGGCGCGCTGAGCCACCGCTGTGTCTACGGGGCTGTGGACC 89
 DB 22 ATACAGGAGCGAGATCTGGCGCGCTGAGCCACCGCT-ATCAGGAGCTGTGGACC 80
 QY 90 AGTTTGAGACCGCAAGACCTCATCTCTGAGCTGTGCTCATCCGAGGAGCTGC 149
 DB 81 AGTTTGAGACCGCAAGACCTCATCTCTGAGCTGTGCTCATCCGAGGAGCTGC 140
 QY 150 TGGACCGCTGTACAGGAGGGGTGTGTCGAGGCGCGAGGTCAAGGTCTATCCAGC 209
 DB 141 TGGACCGCTGTACAGGAGGGGTGTGTCGAGGCGCGAGGTCAAGGTCTATCCAGC 200
 QY 210 AGCTGTGGAGGGGTGTGCTACTCTGCACAGCCATGGCGTTCTCCACCTGGACATAAAGC 269
 DB 201 AGCTGTGGAGGGGTGTGCTACTCTGCACAGCCATGGCGTTCTCCACCTGGACATAAAGC 260
 QY 270 CCTCTAACATCTGTATGGTGCATCTGCCCGGG--AAGACATTAAATCTGCAGCTTTGG 327
 DB 261 CCTCTAACATCTGTATGGTGCATCTGCCCGGGGAAGACCATTCAAATCTGCAGCTTTTG 320
 QY 328 CTTT 331
 DB 321 GATT 324

RESULT 12
 AJ573174
 LOCUS AJ573174 HM3/RH2 Homo sapiens cDNA clone HSPD46382, mRNA sequence.
 DEFINITION
 ACCESSION AJ573174
 VERSION AJ573174.1 GI:33297035
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 371)
 Laveder P., De Pitta C., Vitulo N., Valle G. and Lanfranchi, G.

Tel: 402 762 4366
Fax: 402 762 4390

Email: smithemall.marc.usda.gov

Single pass sequencing. Bases called and trimmed with the -minscore 20
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -minmatch 12 options.

PCR Primers

FORWARD: AGGAACAGCTATGACCAT

BACKWARD: GTTTCCTCAGTCAGGCG

Plate: 28 row: F column: 16

Seq primer: ATTAGTGTCACACTATAG.

Location/Qualifiers

1. 368

/organism="Bos taurus"

/mol_type="mRNA"

/db_xref="taxon:9913"

/tissue_type="pooled"

/lab_host="DH10B"

/clone_lib="MARC 2B0V"

/note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: Sall;

Library made from pooled tissue from testis, thymus,

semimendonsus muscle, longissimus muscle, pancreas,

adrenal, and endometrium."

FEATURES

source

ORIGIN

Query Match 7.6%; Score 229.4; DB 2; Length 368;

Best Local Similarity 81.3%; Pred. No. 1.3e-35;

Matches 282; Conservative 0; Mismatches 56; Indels 9; Gaps 1;

Qy 2588 GCAGCGCTGGAGATCACAGAGAGTCAGAGATGTGGACGGCTCTCTGCAGAGCGTCCC 2647

Db 31 GCCAGCTGGAGATCACAGAGAGCCAGAGATGGAGCGCTCTCGGAAAGCTGCC 90

Qy 2648 GTGGCAGGAAGCGCAAGTGTCTTCGCCGTACAGCAGCTCTTCCACTTCCCTGGGAGG 2707

Db 91 GCGGCGAGGAGCGCAAGTGTCTACCCCTTCGGGCGAGTCTTCCACTTCCCTGGGAGG 150

Qy 2708 CACTGCGCTGGATGACCTGCAGAGCTGGGCTCGTGAGAGTGAAGCTCCCTCGTG 2767

Db 151 CACTCAGCTTCCTGGAGGAGCCACAGAGCTGGGCTACGCCAGAGGTGAAGCGCTCGGTA 210

Qy 2768 GAGCAGATCTCCCGATCTGAAGGCGAGCGGAGGTCTGGAGAGGAGGCGGCCCC 2827

Db 211 GCCAGATCTCCGCTCTGAGGCGAGGCTGAGGTCTGGAGAGGAGTCTCTCC 270

Qy 2828 AGGAAGAGCGAGCGCTTCTTCTTCCTCGGCTCTCAGGTCTGAAGAGCTGGGACGAGCG 2887

Db 271 CGGAAGAGGCGAGCGCTTCTTCTTCCTCC-----GCTTGAAGAGCAAGGACGAGCA 321

Qy 2888 CCGACATCTCTAAGGAGCTCTCAGATGAGACTGTGCTCTGGGCGCA 2934

Db 322 CCGTCTCTTAAGGAGCTCTCGGATGAGAGCGTGTCTTCTGGGCGCA 368

RESULT 14

BY739930

LOCUS

DEFINITION BY739930 RIKEN full-length enriched, pooled tissues, 16 days

embryo, etc. Mus musculus cDNA clone 1920063N21 5', mRNA sequence.

ACCESSION BY739930

VERSION BY739930.1 GI:27163955

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 653)

Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,

Nikaido, I., Oeato, N., Saito, R., Suzuki, H., Yamanaka, I.,

Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,

Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D.P., Bult, C.,

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H.,

Batalov, S., Beisel, K.W., Blake, J.A., Bratt, D., Brusic, V.,

Chothia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A.,
Fletcher, C.F., Forrest, A., Frazer, K.S., Gasterland, T.,
Gariboldi, M., Giesi, C., Godzik, A., Gough, J., Grimmond, S.,
Kutwinski, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A.,
Kawaji, H., Kawaasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A.,
Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Magloct, D.R.,
Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,
Numata, K., Okido, I., Pavan, W.J., Pertea, G., Pesole, G.,
Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S.,
Ravasi, T., Reed, J.C., Reid, J., Ring, B.Z., Ringwald, M.,
Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K.,
Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M.,
Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,
Wells, C., Wilming, L.G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I.,
Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,
Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M.,
Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,
Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y.,
Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K.,
Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S.,
Rogers, J., Birney, E., and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

22354583

12466851

PUBMED

COMMENT

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,

Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,

Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,

Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,

Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,

Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,

Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in Riken.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details

Location/Qualifiers

1. 653

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/clone="1920063N21"

/clone_lib="RIKEN full-length enriched, pooled tissues, 16

days embryo, etc."

/note="pooled tissues: (dev stage=16 days

embryo, tissue type=heart, sex=mix), (dev stage=16 days

embryo, tissue type=kidney, sex=mix), (dev stage=17 days

embryo, tissue type=heart, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

embryo, tissue type=stomach, sex=mix), (dev stage=17 days

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 16:50:48 ; Search time 4029.45 Seconds
(without alignments)
10400.615 Million cell updates/sec

Title: US-10-077-130-6_COPY_13600_14700

Perfect score: 1101

Sequence: 1 cacagcagccacatgtgac.....agcaggaaggcccatgttc 1101

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hcc: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gssi: *
9: gb_gss2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	145	13.2	479	CEI79805	tigr-gss tigr-gss-
2	57.6	5.2	932	CNS0072Q	AL066742 Drosophil
3	56.6	5.1	1218	CL081011	AL081011 CH216-160
4	56	5.1	1242	CD509648	CD509648 CDA97-D11
5	54.8	5.0	925	CNS00891P	AL053013 Drosophil
6	54.8	5.0	975	AG095351	AG095351 Pan trogl
7	53.8	4.9	925	CNS00091P	AL053013 Drosophil
8	53.4	4.9	354	BF353892	BF353892 IL5-HT070
9	52.6	4.8	462	BI976924	BI976924 486128 MA
10	52	4.7	910	CNS00060N	AL065629 Drosophil
11	50.2	4.6	510	BX478907	BX478907 DKF2p686N
12	50	4.5	932	CNS0072Q	AL066742 Drosophil
13	49.6	4.5	784	AG161654	AG161654 Pan trogl
14	49	4.5	500	BF073047	BF073047 219598 MA
15	48.8	4.4	776	CNS0100RY	AL099352 Drosophil
16	48.8	4.4	870	CNS050XKS	AL316153 Tetraodon
17	48.6	4.4	799	BQ435439	BQ435439 AGENCOURT
18	48.6	4.4	971	BQ222099	BQ222099 AGENCOURT
19	48.6	4.4	988	2 BE876681	BE876681 601488245
20	48.6	4.4	1046	BQ643604	BQ643604 AGENCOURT
21	48.6	4.4	1183	CR720606	CR720606 Tetraodon
22	48.6	4.4	1417	AY412116	AY412116 Homo sapi
23	48	4.4	676	CE563102	CE563102 tigr-gss-
24	48	4.4	1033	CG756677	CG756677 P051-4-D0

25	47.6	4.3	666	4	BG786337	BG786337 SEAMC006
26	47.6	4.3	895	2	BE731455	BE731455 601567740
c 27	47.6	4.3	1014	9	CNS042T9	AL314694 Tetraodon
28	47.4	4.3	1116	4	BG810038	BG810038 mgct002x1
c 29	47.2	4.3	808	9	AG030609	AG030609 Pan trogl
30	47.2	4.3	964	9	CNS003WG	AL065254 Drosophil
c 31	47	4.3	1051	5	BUS40105	BUS40105 AGENCOURT
c 32	47	4.3	1101	9	CNS017SY	AL108460 Drosophil
c 33	46.8	4.3	1136	8	CC292742	CC292742 CH261-185
34	46.8	4.3	1286	9	AG448422	AG448422 Mus muscu
35	46.6	4.2	641	9	CE104406	CE104406 tigr-gss-
36	46.6	4.2	730	6	CA379703	CA379703 658844 NC
37	46.6	4.2	1021	5	BQ434910	BQ434910 AGENCOURT
38	46.6	4.2	4219	3	BC033185	BC033185 Homo sapi
39	46.4	4.2	359	5	BQ091517	BQ091517 UNM18H03
40	46.4	4.2	514	2	BE722090	BE722090 189984 MA
41	46.4	4.2	514	2	BE722099	BE722099 189998 MA
42	46.4	4.2	1178	8	CC227732	CC227732 CH261-156
43	46.4	4.2	1425	9	AG441757	AG441757 Mus muscu
c 44	46.2	4.2	613	6	CA222734	CA222734 SCEZF4403
45	46.2	4.2	836	4	BG310024	BG310024 HVSMEC001

ALIGNMENTS

RESULT 1
CEI79805
LOCUS CEI79805 479 bp DNA linear GSS 25-SEP-2003
DEFINITION tigr-gss-dog-17000326766391 Dog Library Canis familiaris genomic,
genomic survey sequence.
ACCESSION CEI79805
VERSION CEI79805.1 GI:35328485
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 479)
AUTHORS Kirchner,B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Rusch,D.B., Halpern,A.L., Levy,S., Remington,K.,
Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirkness@tigr.org
Class: shotgun.
Location/Qualifiers
source
1. 479
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

Query Match 13.2%; Score 145; DB 9; Length 479;
Best Local Similarity 80.9%; Pred.No. 1.1e-21;
Matches 169; Conservative 0; Mismatches 40; Indels 0; Gaps 0;
QY 1 CACAGCAGCCACATGTGACACTGTCTGGGAGCTCCCATCATGATGATGAGCGGTGTG 60
DB 271 CAAAGTGGCCGCTCTGTACCGCTGTCTGCTGGGCGCCAGCAGCATGCTGGCGGTGG 330


```

QY 61 CTCTGTGGCTACCGCTGAGGTGAAGAGGGGGCCACAGGCCAGTGGCGCTGTGCCAC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 331 CTGTGGCGCTATCGGTGGAGATGAGGGCGCCACACCGGGCGAGTGGCAGCTGTGCCAC 390

QY 121 GAGCTGTGTGCTGACCCGAGTGTGTGTGTGATGGCTTGGCCCCCGGGGAGACCTACGC 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 391 GAGCTGTGTGCTGCGGCGCTGAGTGTGATCGTGACGCGCTTGGCCCCCGGGGAGTCTTACGC 450

QY 181 TTCCGTGTGGCAGCTGTGGCCCTGTGGG 209
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 451 TTCCGTGTGGCAGCGGTGGGCCAGTGGG 479

RESULT 2
CNS00720/c 332 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence T7 end of BAC #
DEFINITION BACR14B09 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL066742
VERSION AL066742.1 GI:4945205
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Epiphytoidae; Drosophilidae; Drosophila.
1 (bases 1 to 932)
REFERENCE Genoscope.
AUTHORS Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
FEATURES
source
1..932
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14B09"
/clone_lib="RPCI-98"
/note="end : T7"

ORIGIN
Query Match 5.2%; Score 57.6; DB 9; Length 932;
Best Local Similarity 28.7%; Pred.No.0.038;
Matches 139; Conservative 107; Mismatches 238; Indels 0; Gaps 0;

QY 488 GCCTGTGCTCAGGGCTCCATCTGCTCGGGTGGCCACCTTCAGGTGGCAGTGGCCAG 547
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 893 GCGGSSCGCGCGCGCGCGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 834

QY 548 CTTCTGTGATGAGGCCCTTACGCCAGCTTCCCGCCCGAGCAGCCAGAGGGGTGACC 607
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 833 SGCGSSCGSSGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 774

QY 608 TGCACCTACTGTGGGAGGCCCTTGGCGGAACGTCGATGAGCGGTGAGCCACGCTGG 667
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 773 SCGCGSCGSSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 714

```

```

QY 668 ACTCCATTAGCGAGCTGCCAGAGGAGGACGCGCGCTGTGCAGCGCTTCCACAGAGGCGAG 727
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 713 GSCGSGCGCGGGGSGCGSGGGGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 654

QY 728 AGAGGTGGCACTGATCTCTCTGAAGGCTACTCCACGGCCGATGAGTGGCCGCGCACTG 787
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 653 CGCGCGCGCGSGGCGSSCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCM 594

QY 788 GAGATGCTGACCTCTCACACACAGCTCTGATGATGATGATGATGATGATGATGATGATGAT 847
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 593 VARNAVSVSCCCSCSCMASCCCGCGVSGCGSCSCMSCCCCCGCASCVCAGCCMSAGCM 534

QY 848 TGGTCACCTACCTCAAGAAGGCTGGGAGGCGGAGGCACTCACCACTGGCCAGCAAGTTG 907
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 533 GCGCCMCCCGCCASGSCSCCCVCGBCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 474

QY 908 GGGCCCCAGAGCCCCCTCTGTGAGGCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 967
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 473 VGMGMCMGMGRCVCSAGSGCAGMARMAAACAGAGCGCAGMGCMGCMGCMGCMGCMGCMG 414

QY 968 GCCC 971
Db 413 ACGC 410

RESULT 3
CL081011/c 1218 bp DNA linear GSS 31-DEC-2003
LOCUS CH216-160116_RM4.1 CH216 Xenopus tropicalis genomic clone
DEFINITION CH216-160116, genomic survey sequence.
ACCESSION CL081011
VERSION CL081011.1 GI:40536924
KEYWORDS GSS.
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 1218)
REFERENCE Kremitzki,C., Carter,J., McPherson,J., Warren,W., Graves,T.,
AUTHORS Mardis,E. and Wilson,R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 175000 Std Error: 0.00
Seq primer: RM4 ctcaaggcgatcggtcgagc
Class: BAC ends
High quality sequence start: 79
High quality sequence stop: 190.
FEATURES
source
1..1218
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/strain="Nigerian frog"
/db_xref="taxon:8364"
/clone="CH216-160116"
/sex="male"
/cell_line="Stock 248 F7A2, inbred N7"
/clone_lib="CH216"
/note="Vector: pTARBAC2.1; CHORI-216 Xenopus tropicalis
BAC library"

ORIGIN
Query Match 5.1%; Score 56.6; DB 9; Length 1218;
Best Local Similarity 40.1%; Pred.No.0.065;
Matches 257; Conservative 0; Mismatches 384; Indels 0; Gaps 0;

QY 29 GGGCAGCTCCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
QY 711 CTGTCCACAGGAGGCGACAGAGGAGGTGGC 737
      ||| ||| ||| ||| ||| ||| ||| |||
Db 238 GCGGGGCGGGGGGCGNNNGGGGCGNC 212

RESULT 5
CNS0091P/c
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BACR19D16 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL053013
GSS.
AL053013.1 GI:4934461
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 925)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"

FEATURES
source
1..925
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 5.0%; Score 54.8; DB 9; Length 925;
Best Local Similarity 13.5%; Pred. No. 0.16;
Matches 53; Conservative 177; Mismatches 164; Indels 0; Gaps 0;

QY 515 CGGTGCGCACCTTCAGGTGGCAGTACCAGCCTCTGTGGATGAGGCCCTCAGCCCA 574
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 925 CSBSCSCSCSSBSSCSMTSSNSBSCSCSSBSSSSTSSBSSSSTSSBSSSSTSSBSSS 866

QY 575 GCTTGCCCGCCGAGCACCCAGGAGGTGACCTGACCTACTGTGGAGGCCCTGGCTC 634
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 865 SSGTSSACVKNASSSCCGCCGMBCCMCSSSSSCGGSASARGVKVRSAGAGKRGGS 806

QY 635 GGAAACGTGCGATGAGCCGTCGACCCACGCTGGACTCCATTAGGAGCTGCAGAGGAGG 694
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 805 GGASASHSSSACBSSSSSCASCSWSASSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 746

QY 695 ACGCCGCTCCAGCGCTGCCACAGGAGGCGAGAGGAGGTGGACCTGATCTCTGAAG 754
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 ASAGSVSSASSSSSSSSSSSSSVSCSSVSSMSSCSBSSSSSSASASSSSSSSSSSSS 686

QY 755 GCTACTCCACGCGCGATGAGCTGCGCCGCACTGGAGATGCTGACTCTCTCACACACCA 814
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 TSWSCSCTSMASMAARSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 626
```

```
QY 815 CTGATGATGAGTCCCGGCGCAGGCACCCCTTCCCTGCTCACCCTACCTCAAGAGGCTGGGA 874
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 SMSGGGGSGSVSASSGMSSSSSSGGRSSGGGGGSGSSSSSSSSSSSSSSSGSGSVCS 566

QY 875 GGCAGGCACCTCACACTGCGCCAGCAAGGTTGG 908
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 SGCMCRSCSSAAAAAASCVAAASCGMCMGKSKSG 532

RESULT 6
AG095351
LOCUS
DEFINITION
Pan troglodytes DNA, clone: PTB-096G08.P, genomic survey sequence.
ACCESSION
AG095351
VERSION
AG095351.1 GI:16647174
KEYWORDS
GSS.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of Library PTB
Unpublished
2 (bases 1 to 975)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (02-AUG-2001) Aao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbases@psc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
Clones are derived from the chimpanzee BAC library PTB This BAC end
was generated during the R&D process and may have higher chance of
clone tracking errors.
PRIMERS
Sequencing: -21M13
LIBRARY
Vector : pKS145
R.Site 1 : SacI
R.Site 2 : SacI.
1..975
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-096G08.P"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match 5.0%; Score 54.8; DB 9; Length 975;
Best Local Similarity 46.3%; Pred. No. 0.16;
Matches 182; Conservative 0; Mismatches 212; Indels 0; Gaps 0;

QY 349 GAGGTGTTGAGGTCTCTGGCACAGGGAATGGAGCGCATCCAGCCCGTGGCGGTTTC 408
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88 GGGGGGGGGAAGCCCCCGCCGCAAGAGGCGCCCGCGCGCGCGCGCGCGCGCGCG 147

QY 409 GAGTGTCTCTCCAGAGTTCGCAACAGATGCTGTGTGATCAAGGGCTTCACGGCAGAGAC 468
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 148 GCGGGGGGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 207

QY 469 CAGGGCGAGTACCACTGTGGCTTGGCTCAGGGCTCCATCTGCCCTTGGCGTGCACCTTC 528
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 208 GGGGGCGGGGGGGCCCCCGGGGGGGGGGGCCCCCCCCCGCGCGCGCGCGCGCG 267

QY 529 CAGGTGGCACTGAGCCAGGCCTCTGTGGATGAGGCCCTTCAGGCCCTTGCCTCCCGG 588
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


[illegible]

RESULT 7
CNS0091P
LOCUS
DEFINITION
925 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence TET3 end of BAC #
BAC19016 of RPC1-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION
VERSION
AL053013
AL053013.1 GI:4934461
KEYWORDS
GSS.
SOURCE
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 925)
AUTHORS
Genoscope.

Genoscope:
Direct Submission
TITLE Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```

FEATURES
source
1. 925
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR19D16"
/clone_lib="RPCI-98"
/notes_end="TET3"

```

Query Match	4.9%	Score 53.8	DB 9	Length 925
Best Local Similarity	17.0%	Pred. NO.	0.26	
Matches 48	Conservative 135	Mismatches 99	Indels 1	Gaps 1

	Qy	1	CACAGCAGCCACA	CTGTGACACTCTCTTTGGCAGCTCCCATGACTGATGAGAGCGGTGGT	60
			: : : :	: : : :	
			: : : :	: : : :	
			: : : :	: : : :	
Db		631	CSSSSCSSSSSSTSS	SSTSSTSKSSSGSSSSSYTTKSTASGSGSWAGGGSG	690
			: : : :	: : : :	
			: : : :	: : : :	
			: : : :	: : : :	
Qy		61	CTCTGTGCCTAC	CCCGTGAGAGTGAAGAGGGGCCACAGCCAGTCGCGCTCTGCCAC	120
			: : : :	: : : :	
			: : : :	: : : :	
			: : : :	: : : :	
Db		691	STGSTSSSSSSSST	TSSSVSSGKSSTBSGSSBSSGSSSSSTSSBBSCTTSSSSSS	750
			: : : :	: : : :	
			: : : :	: : : :	
			: : : :	: : : :	
Qy		121	GAGCTGTGGTCT	-GGACCCGAGTGTGTGGATGGCTGCGCCCCGGGAGACACTACCG	179

		: : : : : : : : : : : : : : : : : :	
Db	751	SSYSSSTCSCTCCCSYSSSTSSSSSTGWSSTGSSSSSVGSSSDSTSTCCSCCCY	810
Qy	180	CTTCGGTGTGCACGTGGCGCCCTGTGGGTGCTGGGGAAACCGGTTCACTCGCCCCAGAC	239
Db	811	MCTCCSYMBYCTSTSCGSSSSSGKGYTKCCGGCGSSSTNGMBGTSACSSSSSSC	870
Qy	240	AGTGCGGCTGCAGAGCCAACGAAGCCTGTGCTCTCCCCAGGCC	282
Db	871	SSSSVSSSSKSSASSSSVSSSGSGVSSNSSSSASKSSSGSVS	913

RESULT 8	354 bp	linear	EST 22-NOV-2000
BF353892			
LOCUS	BF353892	354 bp	
DEFINITION	I15-HT0704-290600-108-e09	HT0704	Homo sapiens cDNA, mRNA sequence.
ACCESSION	BF353892		
VERSION	BF353892.1	GI:11312966	
KEYWORDS	EST.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 354)
 Daas Neto, E., Garcia Correa, R., Zago, M.A., Bordin, S., Costa, F.F.,
 Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
 Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H.,
 Brunstein, A., deOliveira, P.S., Bucher, P., Jongenel, C.V.,
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
 Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 199, 4 andar, 01509-010, Sao Paulo-SP,
Brazil

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
<http://www.ludwig.org.br/scripts/gethtml3.pl?tl=IL5&t2=IL5-HT0704-230600-108-e09&t3=2000-06-29&t4=1>
 Seq primer: puc 18 forward
 High quality sequence stop: 341.

```

FEATURES
source
1. .354
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev stage="Adult"
/clone_lib="HT0704"
/notes="Organ: head neck; Vector: puc18; Site 1: Sma1;
Site 2: Sma1; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

```

ORIGIN		low stringency conditions.				
	Query Match	4.9%	Score 53.4;	DB 2;	Length 354;	
	Best Local Similarity	53.6%;	Pred. No. 0.28;			
	Matches 111;	Conservative	0;	Mismatches 96;	Indels 0; Gaps 0;	
Qy	287	CCCCGTGAGACCCGGCAGGTGGCGAGCTGGTGAAGATGTCTCTCTGGAGCTTTGAGGTGGTGG	346			
Db	12	CCCTGGAAGACCAATGTGGTGGCGCCAGGGGAGAGACGTGGAGCTGCGCTGTGAGCTGTCAAC	71			


```

QY 347 CTGAGGCTGGTAGGTCATCTGGCAACAAGGAATGAGCGCATPCAGCCCGGTGGCGGT 406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72 GGC CGGGAACGCCCGCTGCACTGGCTGAAGGACAGGAAGCCATCCCAAGAGCCAGAACT 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 407 TCGAGGTGGTCTCCCAAGGTCGCGACACAGATGCTGGTGATCAAGGCTTCACGGCAGAG 466
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 ATGATGTGGTCTCGGAGGACAGATGGCCATGCTGTGTCATCCCGCGGGCTCGCTCAAGG 191
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 467 ACCAGCGCGGAGTACCACTGTGGCCTGG 493
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 192 ACGCGGCGAGTACACGTGTGAGTGG 218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 9
LOCUS BI976924 462 bp mRNA linear EST 23-OCT-2001
DEFINITION 486128 MARC 2BOV Bos taurus cDNA 5', mRNA sequence.
ACCESSION BI976924
VERSION BI976924.1 GI:16351329
SOURCE EST.
ORGANISM Bos taurus (cow)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 462)
AUTHORS Smith,T.P.L., Grosse,W.M., Freking,W.A., Roberts,A.J., Stone,R.T.,
Casas,E., Wray,J.E., White,J., Cho,J., Fahrenkrug,S.C.,
Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
Chicko-McKown,C.G., Perlea,G., Holt,I., Karamycheva,S., Liang,F.,
Quackenbush,J. and Keele,J.W.
TITLE Sequence evaluation of four pooled-tissue normalized bovine cDNA
libraries and construction of a gene index for cattle
JOURNAL Genome Res. 11 (4), 626-630 (2001)
MEDLINE 21180013
PUBMED 11282978
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCAAGCAGC
Plate: 115 row: 1 column: 9
Seq primer: ATTAGGTGACACTATAG.

```

FEATURES

```

source
1..462
    /organism="Bos taurus"
    /mol_type="mRNA"
    /db_xref="taxon:9913"
    /tissue_type="pooled"
    /lab_host="DH10B"
    /clone_lib="MARC 2BOV"
    /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from testis, thymus,
semitendoneus muscle, longissimus muscle, pancreas,
adrenal, and endometrium."

```

ORIGIN

```

Query Match 4.8%; Score 52.6; DB 4; Length 462;
Best Local Similarity 54.5%; Pred. No. 0.44;
Matches 128; Conservative 0; Mismatches 104; Indels 3; Gaps 1;

QY 12 CACTGTGACATGCTTGGCGAGTCCCATGAGTGATGAGCGGTGCTCTGTGGCTA 71
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 CAGCATCACCTGAGCTGGCTCGGCCCGACACCCAGGACGGGATGAACCGCAGGGCTA 80
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 72 CCGCGTGGAGGTGAAGGAGGGGGCCACAGGCCAGTGGCGGTGTGTCCACAGAG---CTGGT 128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 81 CTTGTGTGAGTCTGCGAGCTGGCCAGCTGTCAGTGGAGCCGTCGCCACGGGGCACAGT 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 129 GCCTGGACCCGAGTGTGTGTGGATGGCTGGCCCCCGGGGAGAGACTTACCGTTCCTGT 188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 141 GAGGGGACCACTACACGCGCAAGGGCTCGGGCCCCGGGAGAGACTTACTTGTTCGGGT 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 189 GCGAGCTGTGGGCGCTGTGGGTGCTGGGGAAACCGGTTTCACTGCCCCAGACAGTG 243
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 201 GACGCGCTGAACGAGGGGGCCCTGGCCAGCCCACTGCTCTGGCTCGGCGGTG 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 10

```

LOCUS CNS0060N 910 bp DNA linear GSS 03-JUN-1999
DEFINITION Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR14J21 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL065629
VERSION AL065629.1 GI:4944698
KEYWORDS GSS.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 910)
AUTHORS Genoscope.
DIRECT SUBMISSION
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammoler in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

```

FEATURES

```

source
1..910
    /organism="Drosophila melanogaster"
    /mol_type="genomic DNA"
    /db_xref="taxon:7227"
    /clone="BACR14J21"
    /clone_lib="RPCI-98"
    /note="end : T7"

```

ORIGIN

```

Query Match 4.7%; Score 52; DB 9; Length 910;
Best Local Similarity 18.4%; Pred. No. 0.66;
Matches 74; Conservative 157; Mismatches 171; Indels 0; Gaps 0;

QY 29 GGGCAGTCCCATGATGAGGGCGGTGCTCTGTGGTACCGCGTGAGGTGAAGG 88
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 465 GGGGGGGKGGGGGGGGGGGGGGGGGGTDTTWTGKSTGGGGGGSSGSSSKG 524
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 89 AGGGGGCCACAGCCAGTGGCGGCTGTGCCACGAGCTGGTGCCTGGACCCGAGTGTGG 148
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 525 GGGGSGSYGGGGGGYGGTGGSGGGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 584
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 149 TGGATGGCTTGGCCCCCGGGGAGACCTTACCGTCTCGTGTGGCAGCTGTGGCCCTGG 208
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 585 CSYSSSSTGSGCGGSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSTSS 644
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db      736  SCCTCCSCCSCGCGSSSGCGSGCGSCGCGSCGCGSCGCGSCGCGSCGCGSCGCGSSGC 795
Qy      345  GGTCTAGAGCTGTGTGAGTTCATCTGGCACAAAGTAATGGAGGCATCCACGCCGCTGGGG 404
Db      796  GSGSGSGCGSSSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGS 855
Qy      405  GTTCGAGGTGGTCTCCAGGTCGCAACAGATGCTGTGTGATCAAGGGCTTCACGGCAGA 464
Db      856  SCGCGSGGCGSGSCGCGSCGCGCGCGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSGSS 915
Qy      465  AGACACAGG 472
Db      916  GCCSCGG 923

```

```

RESULT 13
AG161654
LOCUS   784 bp DNA linear GSS 09-JAN-2002
DEFINITION Pan troglodytes DNA, clone: RP43-027N15.TJ, genomic survey
sequence.
ACCESSION AG161654
VERSION   AG161654.1 GI:16691332
KEYWORDS GSS.
SOURCE    Pan troglodytes (chimpanzee)
ORGANISM  Pan troglodytes
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.

```

```

REFERENCE 1
AUTHORS   Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
          Totoki,Y., Watanabe,H. and Sakaki,Y.
          BAC end sequences of Library RPCI-43
          Unpublished
          Direct Submission
          Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
          and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
          1-7-22 Sushiro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
          (E-mail:chimbese@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
          Tel:81-45-503-9111, Fax:81-45-503-9170)
          Clones are derived from the chimpanzee BAC library RPCI-43 This BAC
          end was generated during the R&D process and may have higher chance
          of clone tracking errors.
          PRIMERS
          Sequencing: TJ
          LIBRARY
          Vector : pBACe3.6
          R.Site 1 : EcoRI
          R.Site 2 : EcoRI.
          Location/Qualifiers
            1..784
              /organism="Pan troglodytes"
              /mol_type="genomic DNA"
              /db_xref="taxon:9598"
              /clone="RP43-027N15.TJ"
              /sex="male"
              /cell_type="lymphocytes"
              /clone_lib="RPCI-43 Chimpanzee Male BAC Library"

```

```

FEATURES
source

```

```

ORIGIN
Query Match      4.5%; Score 49.6; DB 9; Length 784;
Best Local Similarity 45.1%; Pred. No. 2.2;
Matches 296; Conservative 0; Mismatches 354; Indels 7; Gaps 3;

```

```

Qy      69  CTACCGGTGAGGTGAAGAGGGGGCCACAGGCCAGTGGCGGTGTGCCACGAGCTGGT 128
Db      34  CTGCGGGAGGGGGGGGGCGCGCGGGGGGNGCGCGCGGGGGGCCCAAAATCAACCCCCCCC 93
Qy      129  GCCTGGACCGAGTGTGTGTGGTGGGCTCTGGCCCCCGGGGAGACCTACCGCTCCCGTGT 188
Db      94  GCCTGCCCGCCCGCGGGAATAGGGGGGGCGNACGGANACC-CCGGCGGTTTTC 152

```

```

Qy      189  GCGAGCTGTGGGCGCTGTGGGTGTGGGGAACCGGTTTCACTGCCCCACAGACAGTGGGCT 248
Db      153  GCCCGCGCGCGGCGAGCGGGCGCGCGGGGGGCGCGGGGGCGGGGGGGCGGGGGCGG 212
Qy      249  TGCAGAGCCACCGAAGCCTGTGCTCTCCACGCCCTCAGCCCCCTGAGAGCCCGCGAGGTGGC 308
Db      213  GNCGGGGCGCGGGGGGGGGGGGGGTTTCGCGCGCCCCCGCGCGGGGGCGGGGGCGG 272
Qy      309  AGCTGGTGAAGATGTCTCTGTGAGCTTGAGGTGTGGCTGAGGCTGTGTGAGGTCAATCTG 368
Db      273  ACCGCGGGGGGGGGCGCGCGCTGCCCGCGCGCGCGGGGGCGCGTGTGGGGGGGACCC 332
Qy      369  GCACAAGGAATGAGCGCATCCAGCCGCGTGGCGGTTTCGAGGTGGTCTCCAGGGTCG 428
Db      333  GAGCCCCCGGGCGGGCGGGCCGCGCGGGGGGGCGGGCCCCCGGGGGGGGGGGGG 392
Qy      429  GCAACAGATGTGTGTGATCAAGGCTTTCACCGCAGAAAGACAGGGCGAGTACCACTGTGG 488
Db      393  CGGGCGGGGGCGCGGGGGCGCGCGCGCGCGGGGGCGGGGGGGGGGGGGCGGGCGG- 451
Qy      489  CCTGGCTCAGGGCTCCATCTGCGGCTGGCCACCTTCCAGGTGGCACTGAGCCAGC 548
Db      452  -CGGGCGCGGGGCGCGGCTGTGGCGGGGCGCGACCGCGGGCGCGGGCGCGGGCGG 510
Qy      549  CTCTGTGTGATGAGGCCCTCAGCCCCAGCTTCCCGCCCCGAGCGAGCCCGAGGAGGTGACCT 608
Db      511  CGCCCGGGGGGGGGGGGGCGCGCGCGGGGGGCGAGGGAGCGCGCGGGGGGGGGCCCC 570
Qy      609  GCACCTACTGTGGAGAGCCCTTGGCTCGGAAACGTCGATAGAGCCGTGAGCCACGCTGGA 668
Db      571  GCGCG- - - - -GGGGGGGGGGCGCGCGCGGACCGCGCGGGGGGGGGGGGGGGGGCGG 626
Qy      669  CTCCATTTAGCGAGCTGCCAGAGGAGGAGCGCGCTCCAGCGGCTGCCACAGGAGGC 725
Db      627  CCCCCCGCGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 683

```

```

RESULT 14
BF073047

```

```

LOCUS   500 bp mRNA linear EST 25-APR-2001
DEFINITION 219598 MARCH 2BOV Bos taurus cdna 5', mRNA sequence.
ACCESSION BF073047
VERSION   BF073047.1 GI:10866558
KEYWORDS EST.
SOURCE    Bos taurus (cow)
ORGANISM  Bos taurus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
          Bovinae; Bos.

```

```

REFERENCE 1 (bases 1 to 500)
AUTHORS   Smith,T.P.L., Grosse,W.M., Freking,B.A., Roberts,A.J., Stone,R.T.,
          Casas,E., Wray,J.E., White,J., Cho,J., Fahrkrug,S.C.,
          Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A.,
          Chitko-McKown,C.G., Perte,G., Holt,I., Karamycheva,S., Liang,F.,
          Quackenbush,J. and Keefe,J.W.

```

```

TITLE      Sequence evaluation of four pooled-tissue normalized bovine cDNA
          libraries and construction of a gene index for cattle
JOURNAL    Genome Res. 11 (4), 626-630 (2001)
MEDLINE    21180013
PUBMED     11282978

```

```

COMMENT    Contact: Smith TPL
          USDA, ARS, US Meat Animal Research Center
          PO Box 166, Clay Center, NE 68933-0166, USA
          Tel: 402 762 4366
          Fax: 402 762 4390
          Email: smith@email.marc.usda.gov
          Single pass sequencing. Bases called and alt trimmed with phred
          v0.980904.e. Vector identified by cross_match with the -minscore 18
          and -minmatch 12 options.
          PCR Primers
          FORWARD: AGGAACAGCTATGACCAT
          BACKWARD: GTTTCCCGGTCAGCAGC

```


THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:15:10 ; Search time 581.044 Seconds
(without alignments)
10198.307 Million cell updates/sec

Title: US-10-077-130-6_COPY_2000_3000
Perfect score: 1001
Sequence: 1 cgtccaccatccggagtg.....ctgaggcaggggccaatgcc 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004as:*
- 13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	24120	8 ABX11642	Abx11642 Human ser
2	983	98.2	20489	12 ADQ22881	Adq22881 Human ser
3	830.8	83.0	2768	8 ABX71198	Abx71198 Novel hum
4	224	22.4	1251	10 ADC30509	Adc30509 Human nov
5	224	22.4	2488	4 AAH18453	Aah18453 Human cDN
6	221.6	22.1	3956	13 ADR08183	Adr08183 Full leng
7	98.4	9.8	724	4 AAH08421	Aah08421 Human cDN
8	98.4	9.8	2184	4 AAH18256	Aah18256 Human cDN
9	93	9.3	1458	12 ADL24314	Adl24314 AW755252-
10	92.4	9.2	5382	10 ADD14722	Add14722 Human src
11	90.8	9.1	2737	4 AA159251	Aai59251 Human pol
12	90.8	9.1	5378	13 ADR67197	Adr67197 Human bla
13	89.4	8.9	3999	4 AA161037	Aai61037 Human pol
14	88.4	8.8	1645	5 AAF24162	Aaf24162 Human sec
15	88.4	8.8	1948	2 ABV74346	Abv74346 Human il-
16	88.4	8.8	1949	2 ABV74347	Abv74347 Human il-
17	88.4	8.8	2170	6 ABQ54970	Abq54970 Human ova
18	86.2	8.6	2534	11 ADM03261	Agm03261 Human cDN
19	86.2	8.6	7564	8 ABZ24581	Abz24581 Human cel
20	78.6	7.9	14061	6 ABV99363	Abv99363 Human NOV

21	78.6	7.9	14109	6 ABV99362	Abv99362 Human NOV
22	67.8	6.8	2254	3 AAC93403	Aac93403 Human sec
23	67.8	6.8	2254	5 AAF24183	Aaf24183 Human sec
24	66.4	6.6	572	5 AAS42508	Aas42508 Human cDN
25	66.4	6.6	2559	8 ACC46215	Acc46215 Human dit
26	64.8	6.5	716	6 ABK35726	Abk35726 cDNA sequ
27	63.4	6.3	20565	4 AAK89083	Aak89083 Human dig
28	63.4	6.3	20565	4 AAS28908	Aas28908 Human imm
29	63.4	6.3	20565	4 AAL03357	Aal03357 Human rep
30	63.4	6.3	20565	10 ADB31749	Adb31749 Human nov
31	58.2	5.8	1005	6 ABN21414	Abn21414 Human ORF
32	58.2	5.8	2000	8 ADA71938	Ada71938 Rice gene
33	54	5.4	1266	6 ABK35725	Abk35725 cDNA sequ
34	53.4	5.3	2884	10 ADA53433	Ada53433 Human cod
35	53.4	5.3	3575	4 AAH57402	Aah57402 Human eke
36	52.8	5.3	2546	13 ACN39251	Acn39251 Tumour-as
37	52.8	5.3	2695	4 AA158586	Aai58586 Human pol
38	52.8	5.3	2695	5 ADQ98803	Adq98803 DNA encod
39	52.8	5.3	2695	9 ADB48563	Adb48563 Novel hum
40	52.8	5.3	2760	12 ADQ84464	Adq84464 Human tum
41	52.8	5.3	2760	12 ADQ85259	Adq85259 Human tum
42	52.8	5.3	2760	12 ADQ86406	Adq86406 Human tum
43	52.8	5.3	2760	12 ADQ83310	Adq83310 Human tum
44	52.2	5.2	3190	13 ADS09800	Ads09800 Human the
45	52.2	5.2	4953	10 ADE14381	Ade14381 Human int

ALIGNMENTS

RESULT 1
ABX11642
ID ABX11642 standard; cDNA; 24120 BP.
XX
AC ABX11642;
DT 09-MAY-2003 (first entry)
XX
DE Human serine/threonine or protein kinase 12599, cDNA.
XX
KW Human; ss; gene; serine/threonine kinase; protein kinase; 12599;
KW cardiovascular disease; heart failure; myocardial infarction;
KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma;
KW blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
KW haemolytic anaemia; cellular proliferative disorder; cancer;
KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
KW psoriasis; inflammatory bowel disease; rheumatoid arthritis;
KW multiple sclerosis.
XX
OS Homo sapiens.
XX

Key	Location/Qualifiers
5'UTR	1..71
CDS	/*tag= a
	72..23978
	/*tag= b
	/product= "Kinase 12599"
	/note= "This CDS is specifically claimed in claim 2"
3'UTR	23979..24120
	/*tag= c
US2002168742-A1.	
14-NOV-2002.	
15-FEB-2002; 2002US-00077130.	
15-FEB-2001; 2001US-0269201P.	
(MILL-) MILLENNIUM PHARM INC.	
Kapeller-Libermann R, Acton SL;	

DR WPI; 2003-298729/29.
DR P-PSDB; ABG76187.
XX
PT Novel isolated human protein kinase, designated 59079 or 12599
PT polypeptide, useful as diagnostic and therapeutic agents for preventing
PT cardiovascular diseases, proliferative disorders, and protein kinase
PT disorders.
XX
XX
PS Claim 2; Page 58-84; 119pp; English.
PS
CC The invention relates to an isolated human serine/threonine or protein
CC kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule
CC comprising at least 85% identity to the nucleic acids appearing as
CC ABX11641 and ABX11642 or their complement, a naturally occurring variant
CC of the kinases or their fragments. Also included are a non-human host
CC cell containing the nucleic acids, an antibody specific for the proteins,
CC identifying a compound which binds to the kinase (by contacting the
CC kinase or a cell expressing the kinase with a test compound and
CC determining whether the kinase binds to the test compound) and modulating
CC the activity of kinase using the identified compound. The kinases and
CC their encoding nucleic acids are useful as diagnostic and therapeutic
CC agents for preventing a disease or condition associated with an aberrant
CC or unwanted 59079 or 12599 activity in a subject, including
CC cardiovascular diseases such as heart failure, and myocardial infarction;
CC disorders involving blood vessels such as atherosclerosis, and Kaposi's
CC sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia,
CC Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders
CC such as cancer; and protein kinase disorders such as autoimmune
CC disorders, diabetes mellitus, psoriasis, inflammatory bowel disease,
CC rheumatoid arthritis, and multiple sclerosis (many examples of diseases
CC encoding nucleic acids and antibodies are useful in screening assays,
CC detection assays (e.g. forensic biology), and predictive medicine (e.g.
CC diagnostic assays, prognostic assays, and monitoring clinical trials and
CC pharmacogenomics). The kinases and their encoding nucleic acids are
CC useful as query sequences to perform a search against public databases to
CC identify other family members or related sequences. The present sequence
CC encodes the kinase 12599
XX
SQ Sequence 24120 BP; 4719 A; 7265 C; 8278 G; 3858 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 8; Length 24120;
Best Local Similarity 100.0%; Pred. No. 5.1e-195;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGCTACCATCCGGAGGTGCCGCCAGCTGTCACGGGGCGCAGCTGAAGTTGGTGGCCA 2130
Db 2071 CGCTACCATCCGGAGGTGCCGCCAGCTGTCACGGGGCGCAGCTGAAGTTGGTGGCCA 2130

Qy 61 ACCGCATTGAGAGCAGCATCCGGATGGAGTTCGGCTCGGCTGTCAGAGGAGCGCAGCTGC 180
Db 2131 ACCGCATTGAGAGCAGCATCCGGATGGAGTTCGGCTCGGCTGTCAGAGGAGCGCAGCTGC 2130

Qy 121 AGCCGCCAGCCGAGCTGCCGGAGGTGCTGGCTCGGCTGTCAGAGGAGCGCAGCTGC 180
Db 2191 AGCCGCCAGCCGAGCTGCCGGAGGTGCTGGCTCGGCTGTCAGAGGAGCGCAGCTGC 2250

Qy 181 TGGCTGAGCTGTACAGTACAGCTGCGGCTGTGACGTGGCTCAAGGATGGTCGACACACTGT 240
Db 2251 TGGCTGAGCTGTACAGTACAGCTGCGGCTGTGACGTGGCTCAAGGATGGTCGACACACTGT 2310

Qy 241 CCCCAGGCCCCAGTATAGGTGCGGCATCGCCCGGGCGGGGTGCTCTTTGGCGAG 300
Db 2311 CCCCAGGCCCCAGTATAGGTGCGGCATCGCCCGGGCGGGGTGCTCTTTGGCGAG 2370

Qy 301 ATGTGGCCCGGAGCGATGACAGGCTCTTACAGTGCCTCAGCCCGGGGGCGCATCGCCT 360
Db 2371 ATGTGGCCCGGAGCGATGACAGGCTCTTACAGTGCCTCAGCCCGGGGGCGCATCGCCT 2430

Qy 361 ACAGCTCTCCGTCAAGGCTCGCGCTTTCTGCAACAGGACATGGCGGGCAGCTGTG 420
Db 2431 ACCAGCTCTCCGTCAAGGCTCGCGCTTTCTGCAACAGGACATGGCGGGCAGCTGTG 2490

Qy 421 TGGATGCGCTGGCTGGGGGGCCCGCGCAGTTTGTAGTGTGAGACCTCCGAAGCCACAGTCC 480
Db 2491 TGGATGCGCTGGCTGGGGGGCCCGCGCAGTTTGTAGTGTGAGACCTCCGAAGCCACAGTCC 2550

Qy 481 ACGTGCACTGTGTACAAGATGGCATGTGAGCTGGGCCACTCCGGTGAAGCGCTTCTTGACGG 540
Db 2551 ACGTGCACTGTGTACAAGATGGCATGTGAGCTGGGCCACTCCGGTGAAGCGCTTCTTGACGG 2610

Qy 541 AGGATGTTGGGAGCGCGCACCGGCTGGTGGCAGCCACAGTCCACGAGCAGGATGAAGGCA 600
Db 2611 AGGATGTTGGGAGCGCGCACCGGCTGGTGGCAGCCACAGTCCACGAGCAGGATGAAGGCA 2670

Qy 601 CCTACTCTCTCCGCGTGGGGAGGACTCTGTGGAATTTCCGGCTCCGGCTCTCTGAGGCCA 660
Db 2671 CCTACTCTCTCCGCGTGGGGAGGACTCTGTGGAATTTCCGGCTCCGGCTCTCTGAGGCCA 2730

Qy 661 AGTGTGTGTTTGTAAAGGAGCAGCTGGCAGCGAGGAAGCTGCGAGGACAGGACAGGCCA 720
Db 2731 AGTGTGTGTTTGTAAAGGAGCAGCTGGCAGCGAGGAAGCTGCGAGGACAGGACAGGCCA 2790

Qy 721 GTGCCACACTGAGCTGCGAGTGGCCAGGCCAGCCAGAGCGGAGGTGACGTGGTACAAAGGATG 780
Db 2791 GTGCCACACTGAGCTGCGAGTGGCCAGGCCAGCCAGAGCGGAGGTGACGTGGTACAAAGGATG 2850

Qy 781 GGAAGAAGCTGAGCTCCAGCTCGAAAGTGTGCATGGAGGCCACAGGCTGCACGCGCAGGC 840
Db 2851 GGAAGAAGCTGAGCTCCAGCTCGAAAGTGTGCATGGAGGCCACAGGCTGCACGCGCAGGC 2910

Qy 841 TGGTTGTGACAGCAGGACAGGCCAGCGGATGCCCGGGGAGTATAGCTGCGAGGCTGGGGGCC 900
Db 2911 TGGTTGTGACAGCAGGACAGGCCAGCGGATGCCCGGGGAGTATAGCTGCGAGGCTGGGGGCC 2970

Qy 901 AGCGGCTCTCTTCCATCTCGATGTCAAAGAGCCCAAGGTGGTGTGTTGCCAAGGACCCAGG 960
Db 2971 AGCGGCTCTCTTCCATCTCGATGTCAAAGAGCCCAAGGTGGTGTGTTGCCAAGGACCCAGG 3030

Qy 961 TGCCACACAGTGAAGTGCAGGCTGAGGCGCAGGCGGCGGCAATGCC 1001
Db 3031 TGCCACACAGTGAAGTGCAGGCTGAGGCGCAGGCGGCGGCAATGCC 3071

RESULT 2
ADQ22881
ID ADQ22881 standard; DNA; 20489 BP.
XX
AC ADQ22881;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5701.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW ds.
XX
OS Homo sapiens.
XX
FN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
PA (PROT-) PROTEIN DESIGN LABS INC.
XX
PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue

sarcoma.

XX Example 2; SEQ ID NO 5701; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma

CC which comprises obtaining a first soft tissue sample from an individual

CC and a normal soft tissue sample from the same or different individual,

CC determining the expression of a gene in both samples and comparing the

CC expression of the gene in both soft tissue samples, where a higher level

CC of protein expression in the first soft tissue sample indicates the

CC presence of soft tissue sarcoma. The method of the invention has

CC cytosolic applications and may be useful for detecting soft tissue

CC sarcoma, possibly via gene therapy or vaccine production. The nucleic

CC acid sequences may be useful in diagnostic and screening applications.

CC The current sequence is that of a human soft tissue sarcoma-upregulated

CC DNA of the invention. The current sequence is not shown within the

CC specification per se but was submitted in CD format by the inventor.

XX

SQ Sequence 20489 BP; 3917 A; 5815 C; 6983 G; 3225 T; 0 U; 549 Other;

Query Match 98.2%; Score 983; DB 12; Length 20489;

Best Local Similarity 98.2%; Pred. No. 2.4e-191;

Matches 983; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 CGCTCACCATCCGGAGGTGCCCGCCAGCTGCGAGCGGGCGCAGTGAAGTTCGTGGCCA 60

Db 2071 CGCTCACCATCCGGAGGTGCCCGCCAGCTGCGAGCGGGCGCAGTGAAGTTCGTGGCCA 2130

Qy 61 ACGGCATTGAGAGCAGCATCCGGATGAGAGTCCGGCGGGCCCGCCAGCTGACGCCAACA 120

Db 2131 ACGGCATTGAGAGCAGCATCCGGATGAGAGTCCGGCGGGCCCGCCAGCTGACGCCAACA 2190

Qy 121 AGCCGCCAGCCAGCTGCGCGGAGGTGCTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 180

Db 2191 AGCCGCCAGCCAGCTGCGCGGAGGTGCTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 2250

Qy 181 TGGCTGAGCTGTCAGATCAGCTGCGGTGAGTGGTCTGAGTGGTCTGAGTGGTCTGAGTGGT 240

Db 2251 TGGCTGAGCTGTCAGATCAGCTGCGGTGAGTGGTCTGAGTGGTCTGAGTGGTCTGAGTGGT 2310

Qy 241 CCCAGGCCCCAAGTATGAGTGTGAGGATCGCGCGGGCGGGGCTGCTCTTGTGGAG 300

Db 2311 CCCAGGCCCCAAGTATGAGTGTGAGGATCGCGCGGGCGGGGCTGCTCTTGTGGAG 2370

Qy 301 ATGTGGCCCGGAGCATGACAGCTCTACGAGTGGCTCAGCGCGGGGGCGGCATTCGCGCT 360

Db 2371 ATGTGGCCCGGAGCATGACAGCTCTACGAGTGGCTCAGCGCGGGGGCGGCATTCGCGCT 2430

Qy 361 ACCAGCTCTCCGTGCAGAGGCTCCGGCGCTTCTGACAGGACATGCGCGGGAGCTGTG 420

Db 2431 ACCAGCTCTCCGTGCAGAGGCTCCGGCGCTTCTGACAGGACATGCGCGGGAGCTGTG 2490

Qy 421 TGGATGCGCTGGCTGGGGCCCGCGCAGTTGAGTGTGAGACCTCCGAAGCCCAACGTC 480

Db 2491 TGGATGCGCTGGCTGGGGCCCGCGCAGTTGAGTGTGAGACCTCCGAAGCCCAACGTC 2550

Qy 481 ACGTGCACTGTATAAGGATGCGATGAGCTGGGCCACTCCGGTGAAGCTTCTTGCAGG 540

Db 2551 ACGTGCACTGTATAAGGATGCGATGAGCTGGGCCACTCCGGTGAAGCTTCTTGCAGG 2610

Qy 541 AGGATGTGGGACCGGCGCAGCTGTGGGAGCTTCTGAGTGTGAGACCTCCGAAGCCCAACG 600

Db 2611 AGGATGTGGGACCGGCGCAGCTGTGGGAGCTTCTGAGTGTGAGACCTCCGAAGCCCAACG 2670

Qy 601 CCTACTCTCGCGGTGGCGGAGGACTCTGTGGGACTTCCGGCTCCGGCTCTCTGAGGCCA 660

Db 2671 CCTACTCTCGCGGTGGCGGAGGACTCTGTGGGACTTCCGGCTCCGGCTCTCTGAGGCCA 2730

Qy 661 AGGTGGTGTGTTAAGGAGCAGCTGGCAGCGAGGAGCTGACGAGGAGGAGGAGGCA 720

Db 2731 AGGTGGTGTGTTAAGGAGCAGCTGGCAGCGAGGAGCTGACGAGGAGGAGGAGGCA 2790

Qy 721 GTGCCACACTGAGCTGCGAGGTGGCCCGAGGCGCCAGACGAGGAGTGCAGTGTGTAAGGATG 780

Db 2791 GTGCCACACTGAGCTGCGAGGTGGCCCGAGGCGCCAGCGAGTGCAGTGTGTAAGGATG 2850

Qy 781 GGAAGAAGCTCAGCTCCAGCTCGAAGTGTGATGGAGGCCACAGGCTGCACGCGCAGGC 840

Db 2851 GGAAGAAGCTCAGCTCCAGCTCGAAGTGTGATGGAGGCCACAGGCTGCACGCGCAGGC 2910

Qy 841 TGGTTGTGACAGCAGGCGAGGCGGATGCCGGGGAGTATAGCTGCGAGGCTGGGGGCC 900

Db 2911 TGGTTGTGACAGCAGGCGAGGCGGATGCCGGGGAGTATAGCTGCGAGGCTGGGGGCC 2970

Qy 901 AGCGGCTCTCTTCCATCTCGATGTCAAAGAGCCAAAGTGGTGTGTCACCAAGCAGCAGG 960

Db 2971 AGCGGCTCTCTTCCATCTCGATGTCAAAGAGCCAAAGTGGTGTGTCACCAAGCAGCAGG 3030

Qy 961 TGGCACACAGTGTGAGGCTGAGGCTGAGGCGAGGCGCAATGCC 1001

Db 3031 TGGCACACAGTGTGAGGCTGAGGCTGAGGCGAGGCGCAATGCC 3071

RESULT 3

ABX71198

ID ABX71198 standard; cDNA; 2768 BP.

XX AC ABX71198;

XX AC

DT 05-MAR-2003 (first entry)

XX

DE Novel human cDNA sequence #423.

XX

KW Human; gene; ss; nervous system disorder; peripheral neuropathy;

KW Huntington's disease; amyotrophic lateral sclerosis; haemophilia;

KW neurodegenerative disease; Parkinson's disease; Alzheimer's disease;

KW autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;

KW insulin-dependent diabetes mellitus; anaemia; thrombocytopenia; wound;

KW ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;

KW fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;

KW coagulation disorder; cancer; tumour; inflammatory disease; septic shock;

KW Crohn's disease; anaphylaxis; proliferation; chemotactic;

KW differentiation; stem cell growth factor; haematopoiesis; chemokinetic;

KW haemostatic; antiinflammatory; expressed sequence tag; EST.

XX

OS Homo sapiens.

XX

XX WO200281731-A2.

XX

PD 17-OCT-2002.

XX

PF 29-JAN-2002; 2002WO-US001222.

XX

PR 30-JAN-2001; 2001US-00774528.

XX

PA (HYSE-) HYSEQ INC.

PA (GOOD/) GOODRICH R. W.

XX

PI Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

PI Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;

XX

DR WPI; 2003-058563/05.

XX

PT Novel polypeptide useful for treating neurodegenerative diseases, myeloid

PT or lymphoid cell disorders, bone disorders, mechanical and traumatic

PT disorders, coagulation disorders, and inflammatory diseases.

XX

PS Claim 1; Page; 612pp; English.

XX

CC This invention relates to the cDNA sequences encoding an isolated novel

CC human polypeptide. The protein encoded by the nucleic acid of the

CC invention is useful for treating central and peripheral nervous system

CC diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic

CC lateral sclerosis); neurodegenerative diseases (e.g. parkinson's disease,

CC Alzheimer's disease); autoimmune disease (e.g. systemic lupus

CC erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)

CC ; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopaenia)
CC ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,
CC osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head
CC trauma); lung or liver fibrosis; reperfusion injury in various tissues;
CC bacterial, viral or fungal infections; allergic conditions such as
CC allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);
CC cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's
CC disease, anaphylaxis). The protein may be used to inhibit the growth,
CC infection or function of infectious agents such as bacteria, fungi,
CC viruses, or to effect bodily characteristics, biorhythms or circadian
CC cycles of rhythms. The protein may also have
CC proliferation/differentiation, stem cell growth factor, haematopoiesis
CC regulation, immune stimulation or suppressing, chemotactic/chemokinetic,
CC haemostatic and thrombolytic, receptor/ligand, and antiinflammatory
CC activities. The cDNA sequences of the invention are useful for expressing
CC recombinant protein for analysis. The present sequence represents a novel
CC human cDNA sequence of the invention, this sequence is an expressed
CC sequence tag (EST) and was identified using subtractive hybridisation
XX
SQ Sequence 2768 BP; 456 A; 849 C; 995 G; 468 T; 0 U; 0 Other;

Query Match 83.0%; Score 830.8; DB 8; Length 2768;
Best Local Similarity 95.3%; Pred. No. 2.4e-160;
Matches 856; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1 CGCTCACCATCCGGAGGTGCCCGCCAGCCTGCACGGGCGCAGCTGAAGTTCTGTGGCCA 60
Db 1861 CGCTCACCATCCGGAGGTGCCCGCCAGCCTGCACGGGCGCAGCTGAAGTTCTGTGGCCA 1920

QY 61 ACGGCATTTAGAGCAGCATCCGGATGAGGTCCGGGCGGCCCGCCAGGTGACTGCCAACA 120
Db 1921 ACGGCATTTAGAGCAGCATCCGGATGAGGTCCGGGCGGCCCGCCAGGTGACTGCCAACA 1980

QY 121 AGCCGCCAGCCGAGCTCCCGGGAGGTGCTGGCTCGGCTGCACGAGGAGCGCAGCTGC 180
Db 1981 AGCCGCCAGCCGAGCTCCCGGGAGGTGCTGGCTCGGCTGCACGAGGAGCGCAGCTGC 2040

QY 181 TGGCTGAGCTGTACATCAGGCTCGCGCTGTGAGTGGCTGAAGGATGGTGCACACTGT 240
Db 2041 TGGCTGAGCTGTACATCAGGCTCGCGCTGTGAGTGGCTGAAGGATGGTGCACACTGT 2100

QY 241 CCCAGGCCCCAAGTATGAGTGCAGGATCGCGCGGGCGCGGGTCTCTTTGTGGCAG 300
Db 2101 CCCAGGCCCCAAGTATGAGTGCAGGATCGCGCGGGCGCGGGTCTCTTTGTGGCAG 2160

QY 301 ATGTGGCCCGGGAGCATGATCAGGCTCTACGAGTGCCTCAGCGCGGGCGCGCATCCGCT 360
Db 2161 ATGTGGCCCGGGAGCATGATCAGGCTCTACGAGTGCCTCAGCGCGGGCGCGCATCCGCT 2220

QY 361 ACCAGCTCTCCGTGCAAGGCTCTCGCGCTTTCTGCACAAGGACATGGCGGCGCAGCTGTG 420
Db 2221 ACCAGCTCTCCGTGCAAGGCTCTCGCGCTTTCTGCACAAGGACATGGCGGCGCAGCTGTG 2280

QY 421 TGGATGCGGTGGTGGGGCGCGGAGCTTGTGAGTGTGAGACTCCGAGACCCACGCTCC 480
Db 2281 TGGATGCGGTGGTGGGGCGCGGAGCTTGTGAGTGTGAGACTCCGAGACCCACGCTCC 2340

QY 481 ACGTGCACTGTGTACAGGATGGCATGTGAGCTGGGCCACTCCGGTGGCGCTTCTTTCAGG 540
Db 2341 ACGTGCACTGTGTACAGGATGGCATGTGAGCTGGGCCACTCCGGTGGCGCTTCTTTCAGG 2400

QY 541 AGGATGTGGGAGCCGGGACCCGGTGTGTGGAGCCACAGTCAACAGGAGGATGAAGGCA 600
Db 2401 AGGATGTGGGAGCCGGGACCCGGTGTGTGGAGCCACAGTCAACAGGAGGATGAAGGCA 2460

QY 601 CCTACTCTCCGCGGTGGGAGGAGCTGTGGACTTCCGGCTCCGGCTCTGTAGGCCA 660
Db 2461 CCTACTCTCCGCGGTGGGAGGAGCTGTGGACTTCCGGCTCCGGCTCTGTAGGCCA 2520

QY 661 AGGTGGTGTGTTTAAAGGAGCAGCTGGCAGCGAGGAGCTGCAGGAGGAGGAGGAGGCA 720
Db 2521 AGGTGGTGTGTTTAAAGGAGCAGCTGGCAGCGGAGGAGCTGCAGGAGGAGGAGGAGGCA 2580

QY 721 GTGCCACACCTGAGCTGGAGGTGGCCAGGCCAGACGGAGGTGACGTGGGTACAGGATG 780
Db 2581 GCGCCACGCTGAGCTGTGAGGTGGCCAGGCCAGACGGAGGTGACGTGGGTACAGGACG 2640

QY 781 GGAAGAAGCTGAGCTCCAGCTCGAAAGTGTGATGGAGGCCACAGGTGCAACGCCAGGC 840
Db 2641 GGAAGAAGTTCAGCTCCAGCTCGAAAGTGTGATGGAGGCCACAGGTGCAACGCCAGGC 2700

QY 841 TGGTGTGTCAGCAGGCGAGGCCAGCGCATCCGGGGAGTATAGCTGCGAGGCTGGGG 898
Db 2701 TGGTGTGTCAGCAGGCGAGGCCAGCGCATCCGGGGAGTATAGCTGCGAGGCTGGGG 2758

RESULT 4
ADC30509
ID ADC30509 standard; cDNA; 1251 BP.
XX
AC ADC30509;
XX
DT 18-DEC-2003 (first entry)
XX
XX Human novel cDNA sequence, SEQ ID NO:591.
XX
DE Human; diagnostic; drug screening; forensics; gene mapping;
XX biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 17; gene; ss.
XX
OS Homo sapiens.
XX
XX WO2003029271-A2.
XX
XX 10-APR-2003.
XX
XX 24-SEP-2002; 2002WO-US030474.
XX
XX 24-SEP-2001; 2001US-0324631P.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang Ty, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
XX WPI; 2003-371981/35.
DR
DR P-PSDB; ADC31480.
XX
XX New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
XX Claim 1; SEQ ID NO 591; 1185pp; English.
PS
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention

(ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628-ADC33394). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, anaemia, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer. The nucleic acids may also be used as hybridisation probes or primers, and in the recombinant production of a protein. The polypeptides are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human cDNA sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX SEQ Sequence 1251 BP; 329 A; 293 C; 363 G; 266 T; 0 U; 0 Other;

Query Match 22.4%; Score 224; DB 10; Length 1251;
 Best Local Similarity 83.6%; Pred. No. 2.2e-36;
 Matches 254; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 650 CTCTGAGCCCAAGGTGGTGTCTTAAGGACAGCTGGCAGCGAAGCTGCAGGCAGA 709
 DB 891 CCCAGAGCCCAAGGTGGTGTTCGCAAGGAGCAGCCAGCACAGGGAGGTGCAGGCTGA 950

QY 710 GCGAGGAGCCAGTGCACACTGAGCTGGAGGTGGCCAGGCCACAGCGAGGTGACGCTG 769
 DB 951 GCGGGGCGCAGTGCACGCTGAGCTGGAGGTGGCCAGGCCACAGAGGTGACGCTG 1010

QY 770 GTACAGGATGGGAAGAGCTGAGCTCAGCTCGAAGTGTGCATGGAAGGCCACAGGCTG 829
 DB 1011 GTACAAGGATGGGAAGAGCTGAGTTCAGCTCGAAGTGGCGCTGGAGGCGCTGGGCTG 1070

QY 830 CACGCGCAGGCTGGTGTGACGAGCGAGCGAGCCAGCGGATCCCGGGAGTATAGCTGCGA 889
 DB 1071 CACAGGAGGCTGGTGTGACGAGCGAGCGGCGGAGCGAGCGGAGGTGACGCTGCGA 1130

QY 890 GCGTGGGGCGCAGCGGCTCTCTTCCATCTGGATGTCAAAGAGCCCAAGGTGGTGTTCG 949
 DB 1131 GCGAGGGGCTCAGCAGCTCTCTTCCGCTGCAGGTGCGAGTCACTGCTTTGGGATGC 1190

QY 950 CAAG 953
 DB 1191 TGAG 1194

RESULT 5
 AAH18453
 XX AAH18453 standard; cDNA; 2488 BP.
 XX AC AAH18453;
 XX XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:18549.
 XX
 KW Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN EP1074617-A2.
 XX
 XX PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-00116126.
 XX
 XX 29-JUL-1999; 99JP-00248036.
 PR 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.

09-JUN-2000; 2000JP-00241899.
 (HELI-) HELIX RES INST.
 Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 WPI; 2001-318749/34.
 Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs.
 Claim 8; SEQ ID NO 18549; 2537pp + Sequence Listing; English.
 The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification. Where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end sequence and an oligonucleotide comprising a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and in gene therapy. The primers are useful for synthesizing polynucleotides, particularly full-length cDNAs. The primers are also useful for the detection and/or diagnosis of the abnormality of the proteins encoded by the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AAH03166 to AAH13628 and AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632 represent oligonucleotides, all of which are used in the exemplification of the present invention
 Sequence 2488 BP; 561 A; 637 C; 716 G; 574 T; 0 U; 0 Other;
 Query Match 22.4%; Score 224; DB 4; Length 2488;
 Best Local Similarity 83.6%; Pred. No. 2.4e-36;
 Matches 254; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 650 CTCTGAGCCCAAGGTGGTGTTCGTAAGGAGCAGCTGCACGAGGAGCTGCAGGCAGA 709
 DB 891 CCCAGAGCCCAAGGTGGTGTTCGCAAGGAGCAGCCAGCACAGGGAGGTGCAGGCTGA 950

QY 710 GCGAGGAGCCAGTGCACACTGAGCTGGAGGTGGCCAGGCCACAGCGAGGTGACGCTG 769
 DB 951 GCGGGGCGCAGTGCACGCTGAGCTGGAGGTGGCCAGGCCACAGAGGTGACGCTG 1010

QY 770 GTACAAGGATGGGAAGAGCTGAGCTCCAGCTCGAAGTGTGCATGGAAGGCCACAGGCTG 829
 DB 1011 GTACAAGGATGGGAAGAGCTGAGTTCAGCTCGAAGTGGCGCTGGAGGCGCTGGGCTG 1070

QY 830 CACGCGCAGGCTGGTGTTCGACAGGAGCGAGCGGATCCCGGGAGTATAGCTGCGA 889
 DB 1071 CACAGGAGGCTGGTGTTCGACAGGAGCGGCGGAGCGAGCGGAGGTGACGCTGCGA 1130

QY 890 GCGTGGGGCGCAGCGGCTCTCTTCCATCTGGATGTCAAAGAGCCCAAGGTGGTGTTCG 949
 DB 1131 GCGAGGGGCTCAGCAGCTCTCTTCCGCTGCAGGTGCGAGTCACTGCTTTGGGATGC 1190

QY 950 CAAG 953
 DB 1191 TGAG 1194

RESULT 6
 ADR08183/C
 ID ADR08183 standard; cDNA; 3956 BP.

XX ADR08183;
 XX 04-NOV-2004 (first entry)
 XX Full length human cDNA useful for treating neurological disease Seq 1689.
 XX gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;
 KW osteoporosis; neurological disease; Alzheimer's disease;
 KW Parkinson's disease; dementia; short memory; cancer;
 KW sense or motor function; emotional reaction; fear response; panic;
 KW osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
 KW tranquilliser.
 XX Homo sapiens.
 XX EP1447413-A2.
 XX 18-AUG-2004.
 XX 12-FEB-2004; 2004EP-00003145.
 XX 14-FEB-2003; 2003JP-00102207.
 XX 09-MAY-2003; 2003JP-00131452.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;
 XX WPI: 2004-583265/57.
 XX P-PSDB; ADR10139.
 XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX Claim 1; SEQ ID NO 1689; 2686pp; English.
 XX This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNAs. The present invention
 CC describes an immunosay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
 CC cyostatic and tranquilliser activities. This polynucleotide is a full
 CC length human cDNA sequence of the invention. NOTE: This sequence is not
 CC given in the sequence listing of the specification but can be obtained on
 CC CD-ROM from the European Patent Office, Vienna Sub-office.
 XX Sequence 3956 BP; 882 A; 1211 C; 933 G; 930 T; 0 U; 0 Other;
 Query Match 22.1%; Score 221.6; DB 13; Length 3956;
 Best Local Similarity 86.3%; Pred. No. 7.8e-36;
 Matches 245; Conservative 0; Mismatches 39; Indels 0; Gaps 0;
 QY 647 CGTCTGAGCCCAAGGTGGTGTTCAGGACAGCTGGCAGCGAGGAGTGCAGGC 706
 DB 1135 CATCCAGAGCCCAAGGTGGTGTTCAGGAGCAGCGCGCATGCGAGGAGTGCAGGC 1076
 QY 707 AGAGCGAGGAGCCAGTGCACACTGAGCTGGAGTGCGGCCAGGCCAGAGGAGTGAC 766
 DB 1075 GGAGCGGGGGCCAGTGCACGCTGAGCTGGAGTGCGGCCAGGCCAGATGGAGGTGAC 1016
 QY 767 GTGTGTACAGATGGGAGAGCTGAGCTCCAGCTCGAAAGTGTGCTGAGGCGCACAGG 826
 |||||

Db 1015 ATGGTACAAGGACCGGGAAGAGCTGAGCTCCAGCTCAAAAGTGGCGATGGAGGCCAGCGG 956
 QY 827 CTGCACGCGCAGGCTGTTGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 886
 Db 955 CTACACACGAGGAGGCTGGTGTGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 896
 QY 887 CGAGGCTGGGGGCGCAGCGGCTCTCTTCCATCTCGATGTCAAAG 930
 Db 895 TGAGGCGGGGGCGCAACGGCTGTCTTCCGCTGCACGTGGCAG 852
 RESULT 7
 AAH08421
 ID AAH08421 standard; cDNA; 724 BP.
 AC AAH08421;
 XX 26-JUN-2001 (first entry)
 DT Human cDNA clone (5'-primer) SEQ ID NO:5256.
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KW Homo sapiens.
 OS EP1074617-A2.
 XX 07-FEB-2001.
 PD 28-JUL-2000; 2000EP-00116126.
 PF 29-JUL-1999; 99JP-00248036.
 XX 27-AUG-1999; 99JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 XX 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI: 2001-318749/34.
 XX Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX Claim 1; SEQ ID NO 5256; 2537pp + Sequence Listing; English.
 XX The present invention describes primer sets for synthesising 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesising polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

compounds that interact with protein tyrosine kinases and/or protein tyrosine kinase pathways.

Claim 2; SEQ ID NO 116; 139pp; English.

The present invention describes a predictor set comprising a plurality of polynucleotides or polypeptides whose expression pattern is predictive of the response of cells to treatment with a compound that modulates protein tyrosine kinase activity or members of the protein tyrosine kinase pathway. Also described: (1) predicting whether a compound is capable of modulating the activity of cells, comprising obtaining a sample of cells, determining whether the cells express a plurality of markers, and correlating the expression of the markers to the compound's ability to modulate the activity of the cells; (2) a plurality of cell lines for identifying polynucleotides and polypeptides whose expression levels correlate with compound sensitivity or resistance of cells associated with a disease state; and (3) identifying polynucleotides and polypeptides that predict compound sensitivity or resistance of cells associated with a disease state, comprising subjecting the plurality of cell lines to one or more compounds, analysing the expression pattern of a microarray of polynucleotides or polypeptides, and selecting polynucleotides or polypeptides that predict the sensitivity or resistance of cells associated with a disease state by using the expression pattern of the microarray. The polynucleotides and polypeptides have cytostatic activities, and can be used in gene therapy. The polynucleotides and polypeptides are useful in predicting the activity of compounds that interact with protein tyrosine kinases and/or protein tyrosine kinase pathways. These may be used in determining drug sensitivity in patients to allow the development of individualized genetic profiles which aid in treating diseases and disorders (e.g. cancer) based on patient response at a molecular level. The present sequence is used in the exemplification of the present invention.

Sequence 5382 BP; 943 A; 1699 C; 1798 G; 942 T; 0 U; 0 Other;

Query Match 9.2%; Score 92.4; DB 10; Length 5382;
Best Local Similarity 46.7%; Pred. No. 2e-09;
Matches 365; Conservative 0; Mismatches 411; Indels 6;

165	Qy	GAGAGGCGCAGCTGCTGGCTGTGACTGTGTAGATCAGGCTCGGGCTGTGTGACGTGGCTGAAG	224
1703	Db	GAGCGGGTGTGCTGACTTGTGAGCTCTCAAGGGTGGACTTCGCGGCAACTGTGTACAAG	1762
225	Qy	GATGGTTCGACACTGTCCCCAGAGCCCAAGTATGAGGTGCAGGCATCGGCGCGGCGGCGG	284
1763	Db	GATGGGCAGAAGTGTGAGGAGAGCGAGTTCTGTGTGTGAAGATGTGATGGCGCAAAACAC	1822
285	Qy	GTGCTCCTTGTGGAGATGTGGCCCGGGACGATGCAGGCCCTCTACGATGCGTCAAGCCGC	344
1823	Db	CGTCTGATCCTGTGAGGCCAAAGTCCAGACAGTGGCGAGTTTGAGTGCAGGACAGAA	1882
345	Qy	GGGGGCGGCATCGCTACAGACTCTCCGTGTGAAGGCTTCGCGGCGTTCTTGTCACAGGAC	404
1883	Db	GGGGTCTCGGCGCTTCTTTCGCGCTCACTGTGTCCAAGATCTCCCGGTGCACATCGTGGACCCC	1942
405	Qy	ATGCGGGCAGCTGTGTGGATGCGTGGCTGGGGCGCCCGCGCAGTTTGAGTGTGAGACC	464
1943	Db	CGAAGACATGTGTTGTGTGCATGCCATTAACTTCCGAGTGTGTCTGTCTGGCTGTGAGTG	2002
465	Qy	TCCGAAGCCACGTCCA CGTGCACACTGGTCAAGGATGGCATGGAGCTGGGCCCACTCCGGT	524
2003	Db	GACCGAGAGACGCCCTGTGCGTTGTGTCAAGACGCGGCAGGAGGTGGAGGAGA--GT	2059
525	Qy	GAGCGCTTCTGACAGGAGATGTGGGACCGGGCACCGGCTGTGTGGCAGGCCACAGTCACC	584
2060	Db	GACTTCGTGTGTGGAGATGAGGGGCCCACTCGCCGCTGTGTCTGTCGCCGCCCAACCAG	2119
585	Qy	AGCAGGATGAAGSCACTACTCTCTCGCGGTGGGCGAGGACTCTGTGGACTTCCGGCTC	644
2120	Db	CCCTCAGACGGGGCGAGTTTCAGTGGCTGCCTGGAGATGAGTGTGCTTACTTCACTGTC	2179
645	Qy	CGCGTCTCTGAGCCCCAAGTGTTGTTTCTAAGAGCAGCTGGGCACGACGAGAGCTGCAG	704

Db 1822 CGTCTGATCTCTGCTGAGGCCAAAGTCCAGGACAGTGGCGAGTTTGTAGTGCAGGACAGNA 1881
QY 345 GGGGGCCGCGATCGCTACAGCTCTCGTGCAGGCGCTCGCGCGTCTTCTGCACAAGGAC 404
Db 1882 GGGGCTCTGGCGCTTCTTGGCGCTACTGTCCAAGATCTCTCCGTCGACATCGTGCACCCC 1941
QY 405 ATGGCGGGCAGCTGTGTGATCCGCTGGCTGGGGCCCGCGCGCAGTTTGTAGTGTGAGACC 464
Db 1942 CGAGAACATGTCTTCTGTCATGCCATACTCCGAGTGTGTCTGCTGCGCTGTGAGGTG 2001
QY 465 TCCGAAGCCACGTCACGTCGCACTGGTACAAGGATGGCATGGAGCTGGGCCACTCCGGT 524
Db 2002 GACCAGAGGACGCCCTCTGCTGTGGTTGTTACAAAGGCGGCGAGGAGTGGAGGAGA--GT 2058
QY 525 GAGCGCTCTTTCAGAGGATGTGGGAGCGGCGGACCGGCTGGTGGAGGACACAGTCAACC 584
Db 2059 GACTTTCGTGTCTGGAGATGAGGGGCCCATCGCCGCTGGTCTGCCCGCCACCCAG 2118
QY 585 AGGCAGGATGAAGCACCTACTCTCGCGCGTGGCGGAGGACTCTGTGAGCTTCCGGCTC 644
Db 2119 CCCTCAGACGGGGCGAGTTTCAGTGCCTGCTGGAGATGAGTGTGCTTCACTTCTGTC 2178
QY 645 CGCGTCTCTGAGCCCAAGGTGTGTTGCTAAGGAGCAGCTGGCACGCGAGGAAGCTGCAG 704
Db 2179 ACCATCAGACAGCTCTCTCTGTTGATCG--TGTATCCAGCGGCAAGTGTATGTGCA 2235
QY 705 CGAGAGGAGGAGCAGTGCACACTGAGCTGCGAGGTGGCCCGCCAGCGGAGAGGTG 764
Db 2236 GCGGTGCGCTGAGCGTGTGTGCTGACCTGTGAGCTATGCGCGCTTGGGCGAGAGGTG 2295
QY 765 ACGTGTACAGGATGGGAGAGCTGAGCTCCAGCTCGAAGTGTGATGGAGGCCACA 824
Db 2296 CGCTGGACAGGATGGAGAGAGGTGTGTGAGAGCCCGCGCTCTCTCTGCAGAAGAA 2355
QY 825 GCGTCCAGCGCAGCGCTGTTGTGACAGGAGCGGCGGAGTGGCGGAGTATAGC 884
Db 2356 GACACTGTCGCGCGCTGCTGCTGCGCTGTCCAGCTCGAGGACTCCGCGGAGTACTTG 2415
QY 885 TCGAGGCTGGGGGCCAGCGCTCTCTTCCATCTGGATGTCAAGAGCCCAAGTGTGTG 944
Db 2416 TGTGAATTGACGATGAGTGGCGCTCTTCACTGTCAAGTGTGAGAGCCCGCGGAGTGGG 2475
QY 945 TT 946
Db 2476 AT 2477
RESULT 13
AAI61037/c
ID AAI61037 standard; cDNA; 3999 BP.
AC AC
XX AAI61037;
XX
DT 22-OCT-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 5026.
XX
KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia; ss.
XX
OS Homo sapiens.
XX
PN WO200153312-A1.
XX
PD 26-JUL-2001.
XX
PF 26-DEC-2000; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX (HYSB-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR WPI; 2001-442253/47.
DR P-PSDB; AAM41881.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX
PS Claim 1; SEQ ID NO 5026; 10078pp; English.
XX
CC The invention relates to human nucleic acids (AAI57798-AAI61369) and the
CC encoded polypeptides (AAM38642-AAM42213) with nontropic,
CC immunosuppressant and cytostatic activity. The polynucleotides are useful
CC in gene therapy. A composition containing a polypeptide or polynucleotide
CC of the invention may be used to treat diseases of the peripheral nervous
CC system, such as peripheral nervous injuries, peripheral neuropathy and
CC localised neuropathies and central nervous system diseases, such as
CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC utilisation of the activities such as: immune system suppression,
CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
CC part of the printed specification
XX
SQ Sequence 3999 BP; 726 A; 1332 C; 1238 G; 703 T; 0 U; 0 Other;
Query Match 8.9%; Score 89.4; DB 4; Length 3999;
Best Local Similarity 45.8%; Pred. No. 8.1e-09;
Matches 346; Conservative 0; Mismatches 406; Indels 3; Gaps 1;
QY 143 GGAGGTGTGCTCGGCTGCACAGAGGCGCAGCTGTGCTGAGCTGTGATGATGAGGT 202
Db 3168 GTATGTGGCAGCCGCTGCGCTGGAGCGTGTGGTGTGAGCTATGCCGSCCTG 3109
QY 203 TCGCGCTGTGACGTGGGTGAAGGATGGTTCGACACTGTCCCCAGGCCCAAGATGAGGT 262
Db 3108 GGCAGAGGTGCGCTGGACCAAGGATGGAGAGGAGGTGGTGGAGAGCCCGCGCTCTCT 3049
QY 263 GCAGGCATCGCGCGGGCGGGTGTCTTGTGCGAGATGTGCGCGGACGATGACG 322
Db 3048 CGAAGAAGACACTGTCTCCCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2989
QY 323 CCTCTACAGTGCCTGACGCGCGGGGGCGGCATCGCTTACCGAGCTCTCCGTGCAAGGCT 382
Db 2988 CGAGTACTTGTGTGAATTGACGATGAGTGGCGCTCTCTTCACTGTCCACGTCACAGAAC 2929
QY 383 CGCGCGCTTCTGCAAGGACATGCGCGGGCAGCTGTGTGATGCGCGTGGGGGGGCC 442
Db 2928 CCCAGTGGGATCATATACCCCTCGCATGAGGTGACCTTGATGCGCGTGTGCTGCTGCTG 2869
QY 443 GCGCAGCTTGTGAGTGTGAGACTCCGAGCCAGCTCCACGTGCTGCTGCTGCTGCTGCTG 502
Db 2868 TGTGCTGTGATGTGTGAACCTGTCTCGGGAGGATGCCCCCTGTGCGCTGCTGCTGCTGCTG 2809
QY 503 CATGAGCTGGGCCACTCCGCTGAGCGCTTCTTTCAGGAGGATGTGGGGAGCGCGCACCG 562
Db 2808 CTGGAGTGGAGGAGGAGC---GAGGCCCTGCTGCTGAGAGGAGTGGGCCACGCTGCGCG 2752

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 16:50:48 ; Search time 3663.47 Seconds
(without alignments)
10400.615 Million cell updates/sec

Title: US-10-077-130-4_COPY_22500_23500

Perfect score: 1001
Sequence: 1 ggcacgtcagtgacactggc.....ctgtcaagtcgtggacct 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_gsei: *
9: gb_gsei2: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	66.5	929	5	BX409793
2	655.6	65.5	683	2	AW862431
3	570.4	57.0	924	5	BX409792
4	475	47.5	475	4	BI916955
5	453	45.3	477	4	BI916956
6	415.2	41.5	556	4	BI776197
7	405.2	40.5	581	7	CR759117
8	397.4	39.7	771	7	CK601206
9	391.6	39.1	565	7	CR758263
10	374.2	37.4	768	2	BF134040
11	370.8	37.0	811	7	CK599424
12	340.8	34.0	567	4	BI345924
13	324.2	32.4	494	7	CF177913
14	324.2	32.4	653	6	BY739930
15	295	29.5	494	6	CB713368
16	293	29.3	526	5	BO554403
17	274	27.4	397	2	AW479627
18	272.8	27.3	485	6	CB727068
19	255	25.5	328	2	AW482385
20	246	24.6	279	2	BF762805
21	225.2	22.5	427	1	AI006121
22	178.6	17.8	447	2	BB484582
23	171.2	17.1	809	7	CN439177
24	170	17.0	654	9	CR182048

C	25	159.6	15.9	207	7	CR756901
	26	156.4	15.6	3241	3	AK035543
	27	150.2	15.0	597	4	BM488048
	28	148.4	14.8	1030	4	BM552361
C	29	142	14.2	715	5	BM985339
	30	138.2	13.8	621	4	BG792248
	31	135.8	13.6	389	5	BY088395
	32	133	13.3	444	7	CR747491
	33	133	13.3	480	5	EX280322
	34	132	13.2	731	4	BG818748
	35	131.8	13.2	832	7	CK602844
	36	130	13.0	187	2	AW435707
C	37	124.8	12.5	740	7	CK774333
	38	114.8	11.5	902	4	BM018924
C	39	111.4	11.1	773	5	EX435097
	40	96.8	9.7	948	7	CO648953
	41	95	9.5	878	4	BJ729693
C	42	92.2	9.2	551	4	BG001365
	43	92.2	9.2	796	4	BG421646
C	44	92	9.2	462	7	CV334914
C	45	92	9.2	582	7	CV327309

ALIGNMENTS

RESULT 1
BX409793
LOCUS BX409793 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION CS00DF017YK06 5-PRIME, mRNA sequence.
ACCESSION BX409793
VERSION BX409793.2 GI:46926497
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 929)
AUTHORS Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On May 13, 2003 this sequence version replaced gi.30640364.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1025.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0BAF014ZC08_AF01308_2&c=1025.r

FEATURES

Location/Qualifiers
1..929
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF017YK06"
/tissue="TYPE=FETAL BRAIN"
/dev_stages="fetal"
/clone.lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

ORIGIN

Query Match 66.5%; Score 666; DB 5; Length 929;


```

Db      541  GGGCCCAAGATCATCCCTACCAACCCCAAGGACAGACAGAGTGCTGCGGAATACGA 600
Qy      726  GGCCTCAAGGCTGCGCCACCCGACCTGCGCCAGCTGACGAGCCTACCTCAGGCC 785
Db      601  GGCCTCAGAGGCTGCGCCACCCGACCTGCGCCAGCTGACGAGCCTACCTCAGGCC 660
Qy      786  CCGGCACTGTGTCTC 801
Db      661  CC-GCACCTGGTGCTC 675

```

```

RESULT 3
BX409792
LOCUS   BX409792 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
DEFINITION
CS0DF017YK06 5-PRIME, mRNA sequence.
BX409792
VERSION  BX409792.1  GI:30640362
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens

```

```

REFERENCE
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1025.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CS0BAP014ZC08_AF01308_l&c=1025.r

```

```

FEATURES
source
1..924
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DF017YK06"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

ORIGIN

```

Query Match      57.0%; Score 570.4; DB 5; Length 924;
Best Local Similarity 95.2%; Pred. No. 4.9e-119;
Matches 641; Conservative 0; Mismatches 24; Indels 8; Gaps 5;

Qy      1  GGCAGTCAGTGACACTGGCTGCCAGTGTGTCAGCCGACGCTGCCAGGCCACCTGG 60
Db      208 GGCACAGTCAGTGACACTGGCTGCCAGTGTGTCAGCCGACGCTGCCAGGCCACCTGG 267
Qy      61  AGCAAGACGGAGCCCTGGAGAGCAGCAGCGCTGCTCATCTCTGCCACCTCAAG 120
Db      268 AGCAAGACGGAGCCCTGGAGAGCAGCAGCGCTGCTCATCTCTGCCACCTCAAG 327
Qy      121 AACTTCCAGCTTCTGACCATCTCTGCTGGTGGTGTGAGGACCTGGGTGTGTACACCTGC 180
Db      328 AACTTCCAGCTTCTGACCATCTCTGCTGGTGGTGTGAGGACCTGGGTGTGTACACCTGC 387

```

```

Qy      181  AGCCTGAGCAATCGCTGGGACAGTGACACCAAGGCGCTCTCCGGAAGCAGAGCGC 240
Db      388  AGCCTGAGCAATCGCTGGGACAGTGACACCAAGGCGCTCTCCGGAAGCAGAGCGC 447
Qy      241  CCCTCATCTTCGCAATGCGCCGATATCGGGAGAGTGTACCGGATGGGGTGTCTGTGTC 300
Db      448  CCCTCATCTTCGCAATGCGCCGATATCGGGAGAGTGTACCGGATGGGGTGTCTGTGTC 507
Qy      301  TGGAAAGCCGCTGGAAATCTACGGCCCTGTGACCTACATTTGTGCACTAGAGCC 360
Db      508  TGGAAAGCCGCTGGAAATCTACGGCCCTGTGACCTACATTTGTGCACTAGAGCC 567
Qy      361  GGCAGCTGGACCACTGGCTCCGACATCTTTGACTGCTGCTACTCTGACCAAGCTC 420
Db      568  GGCAGCTGGACCACTGGCTCCGACATCTTTGACTGCTGCTACTCTGACCAAGCTC 627
Qy      421  TCCGGGGTGGCACCCTACACCTTCGCGACGGCATGTGTGACCAAGCAGGATGGGTCCC 480
Db      628  TCCGGGGTGGCACCCTACACCTTCGCGACGGCATGTGTGACCAAGCAGGATGGGTCCC 687
Qy      481  TACAGACGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCACTGGCTCTGAGGAG 540
Db      688  TACAGACGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCACTGGCTCTGAGGAG 747
Qy      541  GAGAGCCAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGATTCAGACA 600
Db      748  GAGAGCCAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGATTCAGACA 803
Qy      601  CAGATCCAGAGGGGCGCTTCAGCGTGTGCGGCAATCTGGGAGAAAGCCAGCGGCGG 660
Db      804  CAGATCCAGAGGGGCGCTTCAGCGTGTGCGGCAATCTGGGAGAAAGCCAGCGGCGG 859
Qy      661  CGCTGGCGGCCA 673
Db      860  CGCTGTCCGACA 872

```

RESULT 4

```

BX16955
LOCUS   BX16955
DEFINITION
60317768F1 NIH_MGC_121 Homo sapiens cDNA clone IMAGE:5242025 5',
mRNA sequence.

```

```

ACCESSION BX16955
VERSION    BX16955.1  GI:16180909
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

```

REFERENCE

```

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE   National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11610 row: g column: 18
High quality sequence stop: 475.

```

FEATURES

```

source
1..475
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5242025"
/lab_host="DH10B"
/clone_lib="NIH_MGC_121"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;"

```


TITLE		Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle	
JOURNAL	Genome Res. 11 (4), 626-630 (2001)		
MEDLINE	21180013		
PUBMED	11282978		
COMMENT	Contact: Smith TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4366 Fax: 402 762 4390 Email: smith@email.marc.usda.gov Single pass sequencing, bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options. PCR Primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTITCCCAGTCACGACG Plate: 93 row: C column: 9 Seq primer: ATTATGTCGACACTATAG. Location/Qualifiers source 1..556 /organism="Bos taurus" /mol_type="mRNA" /db_xref="taxon:9913" /tissue_type="pooled" /lab_host="DH10B" /clone_lib="NARC 2BOV" /notes="vector: pCMV SPORT6; Site 1: Noti; Site 2: SalI; Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."		
ORIGIN			
Query Match		41.5%; Score 415.2; DB 4; Length 556;	
Best Local Similarity		85.4%; Pred. No. 9.1e-84;	
Matches 475; Conservative		0; Mismatches 78; Indels 3; Gaps 1;	
QY	118	AAGAACTTCAGACTTCAGCATCCTCGTGTGTGTGTGCTGAGGACCTGGGTGTGTACACC	177
DB	1	AAGAACTTCAGACTTTTGACATCCTGTGTGTGTACTGCTGAGGACCTCGGCATGTATACAG	60
QY	178	TGCAGCGTGAGCAATGCGCTGGGACAGTGACACACGCGCGTCTCCGGAAGGACAGAG	237
DB	61	TGCAGTGTGAGCAATGCGCTGGGACGCGGCCACTACGCCATCTCTCCGAAGGACAGAG	120
QY	238	CGCCCTCATCTTCGCCATGCCGATATCGGGAGGTGTACGGGATGGGTGCTCTGCTG	297
DB	121	CGCCCTCTCTTCCTGCGCACGCGCGGACATTTGGGGAGGTATCGCGGACGGGGTGTCTGCTG	180
QY	298	GTCTGGGAAGCCGTGGAAATCCTACGGGCCCTGTGACCTTACATTTGTGAGTGCAGCTTAGAA	357
DB	181	GTCTGGGAAGCCTGTGGAGTATACGGGCCCTGTGACGTACATTTGTGAGTGCAGCTTGAA	240
QY	358	GGCGGACGTGGACACACATCTGGGCTCCGACATCTTTGACTGTGCTACCTGACGACGCAAG	417
DB	241	GGCGGACGTGGAGACGCTGGCTCGGACGCTCTTTGACTGCTGCTACCTCACAGGAAG	300
QY	418	CTCTCCGGGGTGGCACTTACACCTTCCGACCGGCATGTGTGACGAGGCGAGGATGGGT	477
DB	301	CTTTCTCGGGTGGGGGTACACTTCGACCGGCTGCGGTCAAGAGGGCGGCATGGGC	360
QY	478	CCCTACAGCAGCCCTCGGAGCAAGTCTCTCTGGGAGGGGCCACGACCTCTGGCCTCTGAG	537
DB	361	CCCTACAGCAGCCCTCGGAGCAGGTCTCTCTGGAAGGGCCAGCCGCTTGGCTCTGAA	420
QY	538	GAGGAGAGCCAGG---GGCGGTACGCCAACCCTGTCCCCAGCAAAAGACTTTCGCATTC	594
DB	421	GAGGAGAGTGGTGCACGCGGCGCGGCCGCGGCCCTGCTGCCAGATGACAGACTTCGCTTC	480

Qy	595	CAGACACAGATCCAGAGGGCCGCTTTCAGCGTGTGGGCAATGCTGGAGAGAGCCACG	655
Db	481	CAGACGACAGATGAGAAGGGGCGCTTTCAGCGTGTGGGCAATGCTGGAGAGAGCCACG	540
Qy	655	GGGCGGCGCTGGCCG	670
Db	541	GGGCGCATGCTGGCTG	556
RESULT 7			
CR759117			
LOCUS		581 bp mRNA linear EST 03-SEP-2000	
DEFINITION		CR759117 mdx substracted from control Mus musculus cDNA clone	
ACCESSION		POAB0042B10, mRNA sequence.	
VERSION		CR759117	
KEYWORDS		CR759117.1 GI:51879601	
SOURCE		EST.	
ORGANISM		Mus musculus (house mouse)	
REFERENCE		Mus musculus	
AUTHORS		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 581)	
TITLE		Rouger, K., LeCunff, M., Steenman, M., Potier, M.C., Gibelin, N., Decheene, C.A. and Legr, J.J.	
JOURNAL		Global/temporal gene expression in diaphragm and hindlimb muscles of dystrophin-deficient (mdx) mice	
COMMENT		Am. J. Physiol. Cell Physiol. 283 (C), 773-784 (2002)	
FEATURES		Contact: Genoscope	
source		Genoscope - Centre National de Sequencage	
		2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE	
		Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr	
		Genoscope sequence ID : POAB0042B10F1.	
		Location/Qualifiers	
		1. 581	
		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/db_xref="taxon:10090"	
		/clone="POAB0042B10"	
		/sex="male"	
		/tissue_type="muscle"	
		/clone_lib="mdx substracted from control"	
		/notes="Vector: pCRIT-TOPO; Piecu G., Gros N., Legr J.J., decheene C.A. Mus musculus substracted libraries, detail of cDNA collection: pool of 4 cDNA libraries: C57BL/10mdx diaphragm cDNA substracted from C57BL/10ScSn diaphragm cDNA, C57BL/10ScSn diaphragm cDNA substracted from C57BL/10mdx diaphragm cDNA, C57BL/10mdx hindlimb muscle cDNA substracted from C57BL/10ScSn hindlimb muscle cDNA, C57BL/10ScSn hindlimb muscle cDNA substracted from C57BL/10mdx hindlimb muscle cDNA."	
ORIGIN			
Query Match		40.5%; Score 405.2; DB 7; Length 581;	
Best Local Similarity		82.5%; Pred. No. 1.7e-81;	
Matches 477; Conservative		0; Mismatches 98; Indels 3; Gaps 1	
Qy	1	GGCCAGTCAGTGACACATGGCTGCGAGGTGTCAGCCAGCCAGCGTGCACGCCACCTGG 60	
Db	4	GGCCATCATGTGACATCTGGCTGCGAGGTGTCAGCCAGCCACCTGGCTGACCTGG 63	
Qy	61	AGCAAGACGGAGCCCGCTGGAGAGCAGCGCGTGTCTCATCTTGTGCCACCTCAAG 120	
Db	64	AGCAAGATGGGTCTCTCTGGAGAGCAGCGGCCACCTCTCATCTCTCCACCTGAAG 123	
Qy	121	AACCTCAGCTTCTGACCATCTCTGGTGGTGGCTGAGGACCTGGGTGTGTACACCTGC 180	
Db	124	AACCTCAGCTGTGTGACCATCTCTGGTGGTGGAGAGGAGGATCTGGGGCACATATACCTGC 183	
Qy	181	AGCGTGAGCAATCGCTGGGGACAGTGACACACACGGCGCTCTCCGGAAGCAGAGCGC 240	
Db	184	TGTGTGAGCAACCCGCTGGGGAACAGAGTACACACAGGTGCTCTCCGGAAGCAGAGCGC 243	
Qy	241	CCCTCATCTTTCGCCCATCGCCGGATATCGGGAGGTGTACGCGGATGGGGTGTCTGTGTC 300	


```

244 CCTTCATCTTCTCCAGCCCTGAGGTGGGGAACTATACAAGGATCGGTGCTGCTGTC 303
301 TGAAGCCCGTGAATCTTACGCGCTGTGACCTACATTTGTCAGTGCAGCTAGAGGC 360
304 TAGAAGCCCTGTGAATCTTGTGCGCCCGGTGACCTACATTTGTGAGTGTGTATAGAAG 363
361 GGCAGCTGGACCACTGCGCTCCGACATCTTTGACTGTGCTTACCTGACCAAGCAAGTTC 420
364 GGCAGCTGGACCACTGCGCTCCGACATCTTGTGACTGTGCTTACCTGCTGCTGCTGCTG 423
421 TCCCGGGTGGACCTTACCTTCCGACGATGTGTGAGCAAGGAGCAAGTGTGCTGCTGCTG 480
424 TCGAGGGTGGACCTTATCTTCCGACAGCATGTGTGAGCAAGGAGCAAGTGTGCTGCTGCTG 483
481 TACAGCAGCCCTCGGAGCAAGTCTCTCCGAGGCGCCAGCCACCTGCTGCTGCTGCTGCTG 540
484 TACAGCAGTCTCTGAGCAGGCTCTTGTAGGAGGACCAACCACTGCTGCTGCTGCTGCTG 543
541 GAGAG---CCAGGGCGGTGAGCCCAACCCCTGCCAG 575
544 GAAAGCAACCGGGAGGCGAGCCAGCTTTTTCACAG 581

```

RESULT 8

CK601206
LOCUS
DEFINITION AGENCOURT 17900829 NIH_MGC_234 Rattus norvegicus cDNA clone
IMAGE:7190569 5', mRNA sequence.

ACCESSION CK601206

VERSION CK601206.1 GI:41114391

KEYWORDS EST.

SOURCE Rattus norvegicus (Norway rat)

ORGANISM

Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

1 (bases 1 to 771)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabbs-remail.nih.gov

Tissue Procurement: Howard Jacobs

cDNA Library Preparation: Express Genomics

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LRAM15048 row: d column: 23

High quality sequence stop: 698.

FEATURES

Location/Qualifiers

1..771

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clones="IMAGE:7190569"

/tissue_type="heart, pooled"

/lab_host="DH10B Tona"

/clone_lib="NIH_MGC_234"

/note="Organ: heart; Vector: pExpress-1; Site:1: EcoRV;

Site 2: NotI; RNA obtained from pooled heart tissue from a

mix of male and female animals at 8 wk old. Tissues were

snap-frozen and kept at -80C for two days before RNA

extraction and purification (TRI-reagent method). cDNA was

primed using oligo-dT primer:

5'-pGAGTGTCTAGTCTGAGCGCGGCCCTT-3' and cloned into

the EcoRV/NotI sites of pExpress-1. Size-selection >1.4kb

resulted in an average insert size of 2.2 kb. This primary

library is normalized (non-normalized primary library is
NIH_MGC_233) and was constructed by Express Genomics
(Frederick, MD). Note: this is a NIH_MGC library."

ORIGIN

Query Match 39.7%; Score 397.4; DB 7; Length 771;
Best Local Similarity 82.1%; Pred. No. 1.1e-79;
Matches 469; Conservative 0; Mismatches 101; Indels 1; Gaps 1;

```

QY 1 GGCAGTCAGTGAACATGCGCTGCCAGTGTACGCCAGCCAGCAGTGCACGGCACACTGG 60
    |||||
DB 172 GGCACAATCAGTGACATGCGCTGCCAGTGTGGCCAGCCAACTGCCACAGGCTACCTGG 231
    |||||
QY 61 AGCAAGACGAGAGCCGCCCTCGAGAGCAGCAGCGTGTCTCATCTCTGCCACCTCAAG 120
    |||||
DB 232 AGCAAGATGGGGCCCTTCTGGAGAGCAGCGCCACCTCTCTCATCTCTTCCACCTGAAG 291
    |||||
QY 121 AACTTCCAGCTTCTGACCATCTCTGGTGTGGTCTGAGGACCTGGGTGTGTACACCTGC 180
    |||||
DB 292 AACTTCCAGCTTCTGACCATCTCTGGTGTGGTCTGAGGAGGATCTGGGCACATATACCTGC 351
    |||||
QY 181 AGCGTGAGCAATGCGCTGGGAGCAGTACCAACGAGGGGCTCTCCGGAAGGAGAGGCG 240
    |||||
DB 352 TGTGTGAGCAACCCACTAGGAGACAGCATCACCACAGGTGTCTCTCCGAAAAGCAGAGCGC 411
    |||||
QY 241 CCCTCATCTTCGCCCATGCCCGGATATCGGGAGGTGTACCGGATGGGTGCTGCTGTC 300
    |||||
DB 412 CCTTCATCTTCTCACGCCCGGAGTGGGAACTATACAGGATGAGTGTCTGCTGTC 471
    |||||
QY 301 TGAAGCCCGTGGAACTCTACGCGCTGTGACCTTACATTGTGCAGTGCAGCTAGAAAGGC 360
    |||||
DB 472 TGAAGCCTGTGGAATCCTATGCCCCAGTACCTTACATTGTGCAGTCTGTATAGAAGGA 531
    |||||
QY 361 GGACGCTGGACCACTGGGCTCCGACATCTTTGACTGTGCTGCTGCTGCTGCTGCTGCTG 420
    |||||
DB 532 GGCAGCTGGACAACCCCTGGGCTCTGACATCTCCGACTGTGCTGCTGCTGCTGCTGCTG 591
    |||||
QY 421 TCCCGGGTGGCACCTTACACTTCCGACCGCATGTGTGAGCAAGGAGGAGGAATGGGTCCC 480
    |||||
DB 592 CCCCGGGTGGCATGTATATCTTCCGACAGCATGTGTGAGCAAGGAGGAATGGGCCCC 651
    |||||
QY 481 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGCGCCAGCCAGCAGTGTGAGGAG 540
    |||||
DB 652 TACAGTAGCCCTCAGAACAGGTCTCTCTGGAGGACCCA-CCACCTGGTCTTGTAGGAG 710
    |||||
QY 541 GAGAGCCAGGGCGGTGAGCCCAACCCCTGC 571
    |||||
DB 711 GAAGCAGCCGGGAGACCGCAGCTTCTCC 741
    |||||

```

RESULT 9

CR758263/c

LOCUS

DEFINITION

CR758263 mdx substracted from control Mus musculus cDNA clone

POAB0042B10, mRNA sequence.

ACCESSION

CR758263

VERSION

CR758263.1

GI:51878747

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

Mus musculus

ORGANISM

REFERENCE

AUTHORS

Dechesne, C.A. and Leger, J.J.

Global/temporal gene expression in diaphragm and hindlimb muscles

of dystrophin-deficient (mdx) mice

Am. J. Physiol. Cell Physiol. 283 (C), 773-784 (2002)

CONTACT: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqrefgenoscope.cns.fr, Web: www.genoscope.cns.fr

Genoscope sequence ID: POAB0042B10.1.

SOURCE	ORGANISM	Location/Qualifiers	source
Mus musculus (house mouse)	Mus musculus	1. .565	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		/organism="Mus musculus"	
		/mol_type="mRNA"	
		/db_xref="taxon:10090"	
		/clone="P0A8004ZB10"	
		/sex="male"	
		/tissue_type="muscle"	
		/note="Vector: pCRL1-TOPO; Pietu G., Cros N., Leger J.J., Decheane C.A. Mus musculus substracted libraries, detail of cDNA collection: pool of 4 cDNA libraries: C57BL/10mdx diaphragm cDNA substracted from C57BL/10ScSn diaphragm cDNA, C57BL/10mdx diaphragm cDNA substracted from C57BL/10mdx diaphragm cDNA, C57BL/10mdx hindlimb muscle cDNA substracted from C57BL/10ScSn hindlimb muscle cDNA, C57BL/10ScSn hindlimb muscle cDNA substracted from C57BL/10mdx hindlimb muscle cDNA."	
		39.1%; Score 391.6; DB 7; Length 565;	
		Best Local Similarity 82.6%; Pred. No. 2.1e-78;	
		Matches 461; Conservative 0; Mismatches 94; Indels 3; Gaps 1;	
QY	158	AGGACCTGGGTGTACACTCGACGGTGAAGTAATGCTGGGACAGTACACACGCG 217	
Db	558	AGGATCGGGACATATACCTGCTGTGTGAGCAACCCGCTGGGACAGCATCACACAG 499	
QY	218	GGCTCTCCGAGGAGGAGCGCCCTCATCTTCGCCATGCCGATATCGGGAGGTGT 277	
Db	498	GTGTCTCCGGAAGAGAGCGCCCTTCATCTTCTCCACGCCCTGAGTGGGGAACTAT 439	
QY	278	ACCGGATGGGTGCTGTGTCTGGAAGCCGCTGGAATCTACGGCCCTGTGACCTTACA 337	
Db	438	ACAAGATCGGTGCTGTGTCTAGAGCCCTGTGGAATCTGTGGCCCGGTGACCTACA 379	
QY	338	TTGTGACGTGACGCTTAGAAGCGGCGAGCTGGACACACTGGCTTCGACATCTTTGACT 397	
Db	378	TTGTGACGTGCTGTATAGAAGTAGGCAGCTGGACAACTGGCTTCAGACATCTCTGACT 319	
QY	398	GCTGCTACCTGACCGACAGCTCTCCGGGTGGACCTACACTTCGCGACGGCATGTG 457	
Db	318	GCTGCTACCTGACCGACAGCTCTCGAGGGTGGCATGTATATCTTCGCGACAGCATGTG 259	
QY	458	TCAGCAAGCAGGAATGGGTGCTTACAGCAGCGCCCTCGGAGCAAGTCTCTCGGAGGCG 517	
Db	258	TCAGCAAGCAGGAATGGGCTTACAGCAGCGCCCTCGAAGAGGTCTCTCGGAGGAC 199	
QY	518	CCAGCCACTGGCTCTGAGGAGGA---GAGCAGGCGCGGTGACGCCAAACCCCTGCCCA 574	
Db	198	CCAAACCTGGCTCTGAGGAGGAAGCAGCGGGGGGAGGCGCAGCTTCTCCCA 139	
QY	575	GCACAAAGACCTTCGATTCACACACAGATCCAGAGGGCGCTTACGCTGTGGGCG 634	
Db	138	GCACAAAGACCTTTCCTTCAGATGTCAGATCCGAGGGGGCGCTTACGCTGTGGGCG 79	
QY	635	AATGCTGGAGAGGCGACGCGGGCGCTGGCGGCCAAGATCATCCCTTACCAACCCA 694	
Db	78	AGTGACGAGGAGAGGCAAGTGGCGGGCCCTAGCTGCTAAGATGCTTCCCTTACCAACTG 19	
QY	695	AGGACAAGACAGCAGTGTC 712	
Db	18	AGGACAAGACAGCTGTAC 1	
RESULT 10	BF134040	768 bp mRNA linear EST 24-OCT-2000	
LOCUS	601778492F1 NCI_CGAP_Lu30	Mus musculus cDNA clone IMAGE:4006702 5';	
DEFINITION		mRNA sequence.	
ACCESSION	BF134040		
VERSION	BF134040.1	GI:10973080	
KEYWORDS		EST.	

Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

```

ORIGIN
Query Match      34.0%; Score 340.8; DB 4; Length 567;
Best Local Similarity 75.4%; Pred. No. 7.7e-67;
Matches 490; Conservative 0; Mismatches 77; Indels 83; Gaps 2;

QY 157 GAGGACCTGGGTGTACACCTGACGCTGAGCAATGCGTGGGACAGTGAACCAACG 216
Db 1 GAGGACCTGGGCGTGTACACCTGACGCTGAGCAACGCACTGGGACAGCAACCAACA 60
QY 217 GCGCTCTCCGAGGAGAGCGCCCTCATCTTCGCCATGCGCCGATATGGGAGGTG 276
Db 61 GCCATCTCTCCAGAA-GCAGAGGCTCTCTCTCCCAAGCGCCGAGCATGGTGAATC 119
QY 277 TACGGGATGGGTGCTGCTGTGGAAGCCCGTGGAAATCTCAGGCCCTGTGACCTAC 336
Db 120 TATGCTGATGGGTGCTGCTGTGGAAGCCCTGTGGAGTCATGTGGCCCTGTGACCTAT 179
QY 337 ATTGTGCACTGAGGCTAGAGGCGGAGCTGGACACACACTGGCTTCGACATCTTTGAC 396
Db 180 ATTGTGCACTGAGGCTGAGGCGGAGCTGGACACATTTGGCTTCGACATCTTTGAC 239
QY 397 TGCTGCTACCTGACCAAGCTCTCCGCGGTGGCACTTACACCTTCGCAAGCATGT 456
Db 240 TGCTGCTACCTGACCAAGCTCTCTCGGGTGGGATATACACCTTCGCAAGCATGT 299
QY 457 GTCAAGAGGAGGAGTGGTCTCTACAGAGCCCTCGGACCAAGTCTCTCTGGAGGG 516
Db 300 GTCAAGAGGAGGAGTGGTCTCTACAGAGCCCTCGGACCAAGTCTCTCTGGAGGG 359
QY 517 CCCAGCACTTGGCTCTGTGAGGAGGAGCCAGGGGCGGTGAGCCAAACCCCTGCCAGC 576
Db 360 CCCAGCACTTGGCTCTGTGAGGAGGAGCCAGGGGCGGTGAGCCAAACCCCTGCCAGC 371
QY 577 ACAAGACCTTCGCATTCAGACACAGATCCAGAGGGGCGGCTTCAGGGTGGTGGCNA 636
Db 372 -----GGGGGCGGCTTCAGGTGGTGGGAG 397
QY 637 TGCTGGGAGAGGCGGAGGCGGCGCTGGCGGCAAGATCATCCCTTACCAACCCCAAG 696
Db 398 TGCGGGAGAGGCGGAGGCGGCGCTGGCTGCGCAAGATGCTGCTTACCAACCCCTGAG 457
QY 697 GACAAGACAGCACTGCTGCGGAATACGAGGCGCTTCAAGGGGCTGGGCAACCCGACCTG 756
Db 458 AACAGCACTGCGCTTTCGAAATATGAGGCGCTCAAGGGGCTTGGGCAACCCCTCACCTG 517
QY 757 GCCAGCTGCAGGAGCTTACCTCAGCCCGGCGACCTGTGCTCATCTT 806
Db 518 GCACAGTTGCAGGCTGCTTACCTCAGCCCGGCGACCTGGTCTCATCTT 567

```

```

RESULT 13
CF177813
LOCUS      494 bp mRNA linear EST 28-JUL-2003
DEFINITION 806332 MARC 3P1G Sus scrofa cDNA 5', mRNA sequence.
ACCESSION CF177813
VERSION    CF177813.1 GI:33289589
KEYWORDS  EST.
SOURCE    Sus scrofa (pig)

```

```

ORGANISM  Sus scrofa
           Sus scrofa
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

```

```

REFERENCE 1 (bases 1 to 494)
AUTHORS   Smith,T.P.L., Fraking,B.A., Ford,J.J., Vaillet,J.L., Fox,J.,
           Wise,T.A., Noneman,D.J., Wray,J.E. and Keele,J.W.
TITLE     A second set of porcine ESTs from a pooled-tissue normalized
           library

```

```

JOURNAL   Unpublished (2003)
COMMENT   Contact: Smith TPL
           USDA, ARS, US Meat Animal Research Center
           PO Box 166, Clay Center, NE 68933-0166, USA

```

```

Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: SRG8004 row: M column: 6
Seq primer: GTATACGACTCACTATAGG.

```

FEATURES

```

source
1..494
Location/Qualifiers
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 3P1G"
/note="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including brain, liver, muscle, placenta/endometrium,
ovary, testes, and bone marrow."

```

ORIGIN

```

Query Match      32.4%; Score 324.4; DB 7; Length 494;
Best Local Similarity 80.0%; Pred. No. 4.1e-63;
Matches 395; Conservative 0; Mismatches 96; Indels 3; Gaps 1;

QY 505 CTCCTGGAGGGCCAGCCACTGCGCTCTCAGAGGAGAGCCAGG---GGCGGTACGCC 561
Db 1 CTCCTGGAGGACCCATGCAGCTGCGCTCTGAAGAGGAGAGTGTGCTCATGACCGGCC 60
QY 562 CAACCCCTGCCACACAAAGACCTTCGCATTTCAGACACACATCCAGAGGGGCGCTTC 621
Db 61 CATCCCTCCCATCATGCATACCTTTGCCCTTCACATACACAAATCATTAAGGGGCGCTTC 120
QY 622 AGCGTGTGCGGCAATGCTGGGAGAAAGCCAGCGGGCGGCGCTGGCGGCCCAAGATCATC 681
Db 121 ATTGTGTGAGGCATTCGCGGAGAAATGCCATTTGGCGCGTGTGCTGCCAATATCGTG 180
QY 682 CCCTACCAACCCCAAGGACAAAGACAGTGTGCGCGAATACGAGGCGCTTCAGGGGCTTG 741
Db 181 CGGGGCCACCCCTGAGAACAGGACTGCCGTGCTTCGATTAATATGATGCTCCCTCAAGGGCTTG 240
QY 742 CGCCACCGGACCTGGCGCCAGCTGCACGACGCTACCTCAGCCCCCGGACACCTGTGCTC 801
Db 241 CGCCACCGCTACCTTGGCACATTTGCAGGCTGCCTACCTCCTCCCGGACACCTGTGCTC 300
QY 802 ATCTTGGAGCTGTGCTCTGGCGCGAGCTGTCTCCCTGCGCGAGAGGCGCTTCCTTAC 861
Db 301 ATCTTGGAGTGTGCTCTGGCGCTGATCTGCTCCCTGCGCGAGAGGCGCTTCCTTAC 360
QY 862 TCAGATCCGAGGTGAAGGACTACCTGTGGCAGATGTTGAGTGCCACCCAGTACCTGCAC 921
Db 361 TCATATACATATGTGAAGACTATCTGTGGCAGATGCTGATTGCCACATCATTTACCTGCAC 420
QY 922 AACGAGACATCTCGACCTGCAGCTGAGGTCCGAGAACATGATCATCACCGAATACAC 981
Db 421 GCCAGCGCTATCTTCGATCTAGACCTATGCTCCGAAAAACATGATTGTCCCGAGTACAC 480
QY 982 CTGCTCAAGGTGCT 995
Db 481 CTGCTCAAGATCAT 494

```

```

RESULT 14
BY739930

```

```

LOCUS      653 bp mRNA linear EST 17-DEC-2002
DEFINITION BY739930 RIKEN full-length enriched, pooled tissues, 16 days
           embryo, etc. Mus musculus cDNA clone 1920063N21 5', mRNA sequence.
ACCESSION BY739930
VERSION    BY739930.1 GI:27163955
KEYWORDS  EST.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus

```


THIS PAGE BLANK (USTPO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:02:07 ; Search time 70.3765 Seconds
(without alignments)
9323.380 Million cell updates/sec

Title: US-10-077-130-6_COPY_8500_8900
Perfect score: 401
Sequence: 1 gaagaccagtggtggcgcc.....caccagcccccagcaggagg 401

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818130359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54.8	13.7	3397	4	US-09-949-016-2188
2	54.4	13.6	2768	4	Sequence 2188, Ap
3	51.2	12.8	7218	1	US-09-774-528-426
4	42.4	10.6	809	4	Sequence 14, Appl
5	42.4	10.6	822	4	US-09-949-016-2543
6	42.4	10.6	8260	4	Sequence 4657, Ap
7	42.4	10.6	14809	4	Sequence 14285, A
8	41.4	10.3	4403765	3	Sequence 1032, Ap
9	41.4	10.3	4411529	3	Sequence 2, Appl
10	41.2	10.3	789	3	US-09-103-840A-1
11	40.6	10.1	3575	4	US-09-280-116-114
12	40.2	10.0	289	3	US-09-023-655-1395
13	40.2	10.0	289	3	US-09-007-005-17
14	40	10.0	15644	4	US-09-244-796-17
15	39	9.7	516	4	US-09-902-540-1133
16	39	9.7	6468	4	US-09-902-540-1133
17	39	9.7	22281	4	US-09-902-540-895
18	38.6	9.6	1227	4	US-09-949-016-14778
19	38.6	9.6	2973	4	Sequence 5523, Ap
20	38.6	9.6	3595	4	Sequence 4476, Ap
21	38.6	9.6	9649	4	US-09-902-540-3658
22	38.6	9.6	11254	4	US-09-949-016-15672
23	38.6	9.6	23847	4	Sequence 1040, Ap
24	38.6	9.6	28172	4	US-09-902-540-1177
25	38.4	9.6	1446	4	Sequence 1221, Ap
26	38.4	9.6	2022	4	US-09-902-540-3664
27	38.4	9.6	10593	4	Sequence 4403, Ap
					Sequence 1030, Ap

c 28	38.4	9.6	23847	4	US-09-902-540-1177	Sequence 1177, Ap
c 29	38	9.5	9080	4	US-09-902-540-1963	Sequence 1963, Ap
30	37.8	9.4	15209	4	US-09-902-540-1110	Sequence 1110, Ap
31	37.8	9.4	747	4	US-09-270-767-10494	Sequence 10494, A
32	37.6	9.4	729	4	US-09-902-540-4561	Sequence 4561, Ap
c 33	37.6	9.4	2409	4	US-09-902-540-6842	Sequence 6842, Ap
34	37.6	9.4	3899	4	US-09-902-540-573	Sequence 573, App
35	37.6	9.4	24905	4	US-09-902-540-1225	Sequence 1225, Ap
36	37.4	9.3	1707	4	US-09-949-016-2100	Sequence 2100, Ap
37	37.4	9.3	2032	4	US-09-902-540-8077	Sequence 8077, Ap
38	37.4	9.3	2424	4	US-09-902-540-4231	Sequence 4231, Ap
c 39	37.4	9.3	7811	4	US-09-902-540-824	Sequence 824, App
40	37.4	9.3	10644	4	US-09-902-540-1028	Sequence 1028, Ap
41	37.4	9.3	11465	4	US-09-949-016-13842	Sequence 13842, A
42	37.2	9.3	1905	4	US-09-902-540-2711	Sequence 2711, Ap
43	37.2	9.3	14861	4	US-09-902-540-1127	Sequence 1127, Ap
c 44	37.2	9.3	48974	3	US-08-920-422-17	Sequence 17, Appl
c 45	37	9.2	459	4	US-09-252-991A-9927	Sequence 9927, Ap

ALIGNMENTS

RESULT 1

US-09-949-016-2188
; Sequence 2188, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2188
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2188

Query Match 13.7%; Score 54.8; DB 4; Length 3397;
Best Local Similarity 47.2%; Pred. No. 0.00016;
Matches 167; Conservative 0; Mismatches 187; Indels 0; Gaps 0;

QY	12	GGTGGCGCCAGGGAGCGTGGAGCTGGCTGGTGGAGCTGTACCGGGCGGGAACCCCGT	71
DB	354	GGTGAGCAAGGCCCAAGATCCCGTGGAACTGGTGACCTGACCTGAGGT	413
QY	72	GCACCTGGCTGAAGGACAGAGGCGCATCCGCAAGAGCCAGAGTATGATGTGCTGCGA	131
DB	414	CAATGGCTCAGATATGCGCAGAGATCCAGATGAGCGGCGAGTACATCTTTGAGTCCAT	473
QY	132	GGGCACGATGCGCATGTGCTCATCCGGGGGCTTCGCTCAAGAGCGCGGGGAGTACAC	191
DB	474	CGGTGCCAAGCGTACCTTGACCATCAGCCAGTGTCTCATTTGGCGGACGACGCGTACCA	533
QY	192	GTGTGAGTGGAGGCTTCCAAAGACAGCAGCGCTCCATGTGGAGAAAGCAACTG	251
DB	534	GTGCGTGGTGGTGGCGAGAGTGTAGCAGCGAGCTCTTTGTGAAGAGCCCCCTGTGCT	593
QY	252	CTTACACAGAGCGTGTACCAATCTGAGTGGAGAGAAAGCAGAGCTGTGTCACGTG	311
DB	594	CATCAGCGCCCTTTGAGGAGCACCTGCTGTGATGTGGGCGAGCGGTGAGTTGAGTG	653
QY	312	CAAGCGGAGCACCCCGCGGCCACAGTACCTTGGCGCAAGGCGCTCTTTGGAGCT	365


```
Db 654 TGAAGTATCGAGGAGGGGGCGCAAGTCAAAATGGCTGAAGGACGGGGTGGAGCT 707
|||||
RESULT 2
US-09-774-528-426
; Sequence 426, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Kyle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 426
; LENGTH: 2768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)..(2738)
US-09-774-528-426

Query Match 13.6%; Score 54.4; DB 4; Length 2768;
Best Local Similarity 48.9%; Pred. No. 0.00019;
Matches 179; Conservative 0; Mismatches 181; Indels 6; Gaps 1;

Qy 25 GAGGACGTGGAGCTGCGCTGTGAGCTGTCAAGGCGGGGAAAGCCCGTGCATGCGCTGAAG 84
|||||
Db 2025 GAGGAGGCGCAGCTGCTGCGCTGAGCTGTCAAGTCAAGCTGCGGCTGTGACGTGGCTGAAG 2084
|||||

Qy 85 GACAGGAAGGCATCCGCAAGAGCCAGCAAGTATGATGTGTTCTGCGAGGGCACGATGCC 144
|||||
Db 2085 GATGGTCGCACACTGTCCCGAGGCCCCAAGTATGAGGTGCAGGCATCGCGCGGGCGG 2144
|||||

Qy 145 ATGCTGTGTCATCCGCGGGGCTCGCTCAAGGACGGCGGAGTACACGTGTGAGTGGAG 204
|||||
Db 2145 GTGCTCTTGTGGAGATGTGGCCCGGAGCATGACGAGCCTCTACGATGCGTCAAGCCG 2204
|||||

Qy 205 GCTTCCAGAGCAGACCGCAGCTCCATGTGGAAGAAAGCAAACTGCTTCACAGAGGAG 264
|||||
Db 2205 GGGGCGCGCATCGCTACAGCTCTCCGTGCAAGGCTCGCGCTTTCTGCACAAGGAC 2264
|||||

Qy 265 CTG-----ACCAATCTCAGGTGGAGAGAAAGCACAGCTGTGTTACGTGCAGACG 318
|||||
Db 2265 ATGGCGGCGCAGCTGTGTGGATGCCGTGCTGGGGCCCGGCGAGTTTGAAGTGTGAGACC 2324
|||||

Qy 319 GAGCACCCCGGGCCACAGTACCTGGCGCAAGGGCCTTTTGGAGCTACGGGCTCAGGG 378
|||||
Db 2325 TCCGAGCCACGTCACAGTGCATGTTGTACAGGATGGCATGGAGCTGGGCCATCCGGT 2384
|||||

Qy 379 AAGCAC 384
|||||
Db 2385 GAGCGC 2390
|||||

RESULT 3
```

```
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-File
US-08-232-463-14

Query Match 12.8%; Score 51.2; DB 1; Length 7218;
Best Local Similarity 4.3%; Pred. No. 0.0015;
Matches 17; Conservative 219; Mismatches 162; Indels 0; Gaps 0;

Qy 1 GAAGACCAGTGGTGGCGCCAGGAGGAGCTGGAGCTGCGCTGTGAGCTGTCAAGGGCG 60
|||||
Db 1429 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1370
|||||

Qy 61 GGAAGCGCCGTGCACTGGCTGAAGCAGCAGGAAGCCCATCCGAAGAGCAGCAAGTATGAT 120
|||||
Db 1369 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1310
|||||

Qy 121 GTGCTGTGCGAGGCGACATGCGCCATGCTGGTGTATCCGCGGGGCTGCTCAAGCAGCG 180
|||||
Db 1309 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1250
|||||

Qy 181 GGCAGGTACAGTGTGAGGTGGAGGCTTCCAAGAGCAGCAGCAGCTCCATGTGGAAGAA 240
|||||
Db 1249 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1190
|||||

Qy 241 AAAGCAAACCTTTCACAGAGGAGCTGACCAATCTGCAGGTGGAGGAGAAAGGACAGCT 300
|||||
Db 1189 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1130
|||||

Qy 301 GTGTTACGTGCAAGAGCGGAGCACCCTGGCGCCACAGTGCAGCTGGCGCAAGGGCCTCTTG 360
|||||
```


Qy 193 TGTGAGGTGGAGGCTTCCAGAGACACAGCCAGCCTCCATGTGGAGAAAAAGCAAACTGC 252
Db 3414 GTCAATTGCTGCTCTCCAAAGTCAAGTTCAGCTTCAAGGTCATAGAGGCAGGTAAGATC 3473
Qy 253 TTCA 256
Db 3474 CTGA 3477

RESULT 7
US-09-902-540-1032
; Sequence 1032, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1032
; LENGTH: 14809
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1032

Query Match 10.6%; Score 42.4; DB 4; Length 14809;
Best Local Similarity 48.1%; Pred. No. 0.31;
Matches 151; Conservative 0; Mismatches 161; Indels 2; Gaps 1;
Qy 49 CTGTACGGGCGGGAACCCCGTGCATCTGGTGAAGGACAGAAAGGCAATCCGCAAGAGC 108
Db 2485 CTACGACGTGCGGAACGCC- -GGAGGGCGGTGGGCTCTTACCCGCTCATCCTC 2542
Qy 109 CAGAAGTATGATGGTCTGCGAGGACAGTGGCCATCTGGTTCATCCGGGGGCTCG 168
Db 2543 CACGCGAGGCGATGCGCGCGACGCGGCGACGCCCGCATGGACCTAGTGGCGCG 2602
Qy 169 CTCAGGACGCGGGCGAGTACAGCTGTGAGGTGAGGCTTCCAGAGACACAGCAGCCTC 228
Db 2603 GCGCGCTTCGACGGCTTCTTGAGTTGCGCTTGAAGCGTGGACATCGCGGGGCTCG 2662
Qy 229 CATGTGGAGAAAAAGCAAACTGCTTCACAGAGAGCTGACCAATCTTCAGCGTGGAGAG 288
Db 2663 CTGCTGTGGAGGAGCGGGCGGCTCATCGTTCACATCTCGGCGAGCCCTTCGACGTG 2722
Qy 289 AAAGGCACAGTGTCTTCACTGCAAGACGGACGCCCGCGGCACAGTGCACCTGGCGC 348
Db 2723 CTGCGCGCGACGTGATGTCCTCGCGCGCGGCTGGCGCGGCGCTCTGTCCAGGCG 2782
Qy 349 AAGGCGCTCTTGA 362
Db 2783 AAGCGCTTCGTGA 2796

RESULT 8
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A

; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Query Match 10.3%; Score 41.4; DB 3; Length 4403765;
Best Local Similarity 56.1%; Pred. No. 1.9;
Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
Qy 28 GAGCTGAGCTGCGCTGTGAGCTGTACCGGCGGGAACGCCGTGCACTGGCTGAAGGAC 87
Db 1194752 GAGCGCGAGCGCGCTCTGGTTTACAGGGTGTGCGGCGCGACGACTTGTCTGACCGAA 1194693
Qy 88 AGGAAGGCCATCCGCAAGAGCCAGAGTATGATGTGTCTGCGAGGGCACATGGCCATG 147
Db 1194692 GCCAGGCGCCACTGCGCACGACCATTTCCAGATGTGGCCTTCGGCGCGCGATGCCCAAG 1194633
Qy 148 CTGCTCATCCGCGGGGCT 166
Db 1194632 GAGCGCGTCAACCGGGCTT 1194614

RESULT 9
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Query Match 10.3%; Score 41.4; DB 3; Length 4411529;
Best Local Similarity 56.1%; Pred. No. 1.9;
Matches 78; Conservative 0; Mismatches 61; Indels 0; Gaps 0;
Qy 28 GAGCTGAGCTGCGCTGTGAGCTGTACCGGCGGGAACGCCGTGCACTGGCTGAAGGAC 87
Db 1194537 GAGCGCGAGCGCGCGGTCTGGTTTACAGGGTGTGCGGCGCGACGACTTGTCTGACCGAA 1194478
Qy 88 AGGAAGGCCATCCGCAAGAGCCAGAGTATGATGTGTCTGCGAGGGCACATGGCCATG 147
Db 1194477 GCCAGGCGCCACTGCGCACGACCATTTCCAGATGTGGCCTTCGGCGCGCGATGCCCAAG 1194418
Qy 148 CTGCTCATCCGCGGGGCT 166
Db 1194417 GAGCGCGTCAACCGGGCTT 1194399

RESULT 10
US-09-280-116-114/c
; Sequence 114, Application US/09280116A

RESULT 11
US-09-023-655-1395
; Sequence 1395, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: COMPOSITION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREMITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US

```

RESULT 12
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558
; GENERAL INFORMATION:
; APPLICANT: Szostak, Jack W.
; APPLICANT: Roberts, Richard W.
; APPLICANT: Liu, Rihe
; TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
; TITLE OF INVENTION: FUSIONS
; FILE REFERENCE: 00786/350003
; CURRENT APPLICATION NUMBER: US/09/007,005B
; CURRENT FILING DATE: 1998-01-14
; EARLIER APPLICATION NUMBER: 60/035,963
; EARLIER FILING DATE: 1997-01-27
; EARLIER APPLICATION NUMBER: 60/064,491
; EARLIER FILING DATE: 1997-11-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 289
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translation template
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(289)
; OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

```

	Query Match	10.0%	Score 40.2;	DB 3;	Length 289;
	Best Local Similarity	8.2%;	Pred. No. 0.47;		
	Matches	19; Conservative	96; Mismatches	116; Indels	0; Gaps 0;
Qy	18	GCCAGGGGAGCGTGGAGCTGCCTGTGACGTGTGCAGGCGGGGAACCCCGTGCACTG	77		
	:	: :	: :	: :	: :
	:	: :	: :	: :	: :
Db	47	RCPARURGRNRRNRNRRNRNRRNRNRRNRNRRNRNRRNRNRRNRNRRNRNRRNRNRRNRN	106		
	:	: :	: :	: :	: :
	:	: :	: :	: :	: :

Db	262	GGCGAGTTGCTGGCCGAGTTCAACACCCGCGAGGACCCCGAGCGGCTCCACGTGGACCGA	321
Oy	241	AAAGCAAACCTGCTTCACACAGAGGAGCTGACCA	271
Db	322	GAACCATCTACCTGCTGGTGGAGCGCACCA	352

Search completed: March 21, 2005, 14:39:18
Job time : 85.3765 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:02:07 ; Search time 526.693 Seconds
(without alignments)
9323.380 Million cell updates/sec

Title: US-10-077-130-6_COPY_19500_22500
Perfect score: 3001
Sequence: 1 cctacggagcagaactcggg.....ccacctggagcaaacgga 3001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2997.8	99.9	5207	4	US-09-858-664A-1
2	2997.8	99.9	5207	4	US-10-274-978-1
3	2997.8	99.9	5207	4	US-10-274-978-3
4	2997.8	99.9	5207	4	US-10-697-263-1
5	2997.8	99.9	5207	4	US-10-697-263-3
6	172	5.7	1429	2	US-09-159-385-4
7	172	5.7	1429	3	US-09-186-277-4
8	170.2	5.7	2157	4	US-09-949-016-1966
9	168.6	5.6	2105	4	US-09-949-016-195
10	168.6	5.6	2132	2	US-09-159-385-3
11	168.6	5.6	2132	3	US-09-186-277-3
12	148.6	5.0	1847	4	US-09-949-016-5449
13	148.6	5.0	5926	4	US-09-917-254-41
14	143	4.8	1788	4	US-09-797-039-9
15	143	4.8	2046	4	US-09-797-039-7
16	128.4	4.3	1584	4	US-09-799-451-205
17	122.2	4.1	1282	2	US-08-878-989-12
18	122.2	4.1	1282	3	US-09-272-796-12
19	122.2	4.1	1282	4	US-09-016-434-953
20	122.2	4.1	1349	1	US-07-951-715A-20
21	122.2	4.1	1349	2	US-08-459-448A-20
22	122.2	4.1	1349	3	US-08-459-595A-20
23	122.2	4.1	1349	3	US-08-459-504B-20
24	122.2	4.1	1349	3	US-08-459-444-20
25	122.2	4.1	1349	3	US-09-547-422-20
26	122.2	4.1	1349	3	US-09-968-462-20
27	121.2	4.0	8865	4	US-09-949-016-4238

28	121.2	4.0	8906	2	US-08-826-267-1	Sequence 1, Appli
29	118.6	4.0	1442	4	US-09-949-016-1590	Sequence 1590, Ap
30	118.6	4.0	1442	4	US-09-949-016-1591	Sequence 1591, Ap
31	118.6	4.0	1480	4	US-09-016-434-1454	Sequence 1454, Ap
32	116.2	3.9	1458	4	US-09-230-896C-5	Sequence 5, Appli
33	112.4	3.7	1074	4	US-09-733-388-3	Sequence 3, Appli
34	112.4	3.7	1074	4	US-10-446-175-3	Sequence 3, Appli
35	112.4	3.7	1158	4	US-09-733-388-1	Sequence 1, Appli
36	112.4	3.7	1158	4	US-10-446-175-1	Sequence 1, Appli
37	112.4	3.7	1671	4	US-09-733-388-5	Sequence 5, Appli
38	112.4	3.7	1671	4	US-10-446-175-5	Sequence 5, Appli
39	112.4	3.7	1733	4	US-09-620-312D-526	Sequence 526, App
40	101.4	3.4	1694	4	US-09-579-664B-3	Sequence 3, Appli
41	101.4	3.4	1694	4	US-10-355-975A-3	Sequence 3, Appli
42	92	3.1	15376	4	US-09-949-016-11937	Sequence 11937, A
43	92	3.1	16621	4	US-09-949-016-13708	Sequence 13708, A
44	90.4	3.0	5228	3	US-09-428-711A-15	Sequence 15, Appli
45	89.6	3.0	2447	4	US-09-960-643-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-858-664A-1
; Sequence 1, Application US/09858664A
; Patent No. 6482624
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP
; CURRENT APPLICATION NUMBER: US/0958,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-858-664A-1

Query Match	99.9%	Score	2997.8;	DB	4;	Length	5207;		
Best Local Similarity	99.9%	Pred. No.	0;						
Matches	2999;	Conservative	0;	Mismatches	2;	Indels	0;	Gaps	0;
Qy	1	CCTACGGAGCAGAACTCGGGCCCGGATACAGGAGCGAGACATCTCTGGCGCGCTGAG	60						
Db	699	CCTACGGAGCAGAACTCGGGCCCGGATACAGGAGCGAGACATCTCTGGCGCGCTGAG	758						
Qy	61	CCACCCGCTGGTCAACGGGGCTGTGGACCAAGTTTGAGACCCGCAAGACCCCTCATCTCTCAT	120						
Db	759	CCACCCGCTGGTCAACGGGGCTGTGGACCAAGTTTGAGACCCGCAAGACCCCTCATCTCTCAT	818						
Qy	121	CCTGGAGCTGTGTCTATCCGAGAGCTGTCTGGACCGCTCTACAGGAAGGGCGTGTGTGAC	180						
Db	819	CCTGGAGCTGTGTCTATCCGAGAGCTGTCTGGACCGCTCTACAGGAAGGGCGTGTGTGAC	878						
Qy	181	GGAGGCGGAGGTCAAGGTCTACATCCAGCAGCTGTGTGGAGGGGTGCACCTACCTGCACAG	240						
Db	879	GGAGGCGGAGGTCAAGGTCTACATCCAGCAGCTGTGTGGAGGGGTGCACCTACCTGCACAG	938						
Qy	241	CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAACTCCTGATGTGTGATCTCTGCCCG	300						
Db	939	CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAACTCCTGATGTGTGATCTCTGCCCG	998						
Qy	301	GGAGACATTTAAATCTGCGACTTTTGGCTTTGGCCAGAAATCAATCAACCCAGCAGAGCTGCA	360						
Db	999	GGAGACATTTAAATCTGCGACTTTTGGCTTTGGCCAGAAATCAATCAACCCAGCAGAGCTGCA	1058						

Qy	361	GTTCAGCCAGTACGGCTCCCTCGATTTCGTCTCCCCCGAGATCATCCAGCAAGAACCCCTGT	420
Db	1059	GTTCAGCCAGTACGGCTCCCTCGATTTCGTCTCCCCCGAGATCATCCAGCAAGAACCCCTGT	1118
Qy	421	GAGCGAAGCCCTCCGACATTTTGGGCCCATGGGTGTATCTCTTACTTCAGCTGACCTGACTGCTC	480
Db	1119	GAGCGAAGCCCTCCGACATTTTGGGCCCATGGGTGTATCTCTTACTTCAGCTGACCTGACTGCTC	1178
Qy	481	ATCCCCCATTTTCCGCGCGAGAGTGACCGTGCCACCCCTCCTGAAACGTCTTGAGAGGGCGCGGT	540
Db	1179	ATCCCCCATTTTCCGCGCGAGAGTGACCGTGCCACCCCTCCTGAAACGTCTTGAGAGGGCGCGGT	1238
Qy	541	GTCATGGAGCAGCCCATATGGTGCCCACTCAGCGAAGAGCCCAAGAATTCATCAAGGC	600
Db	1239	GTCATGGAGCAGCCCATATGGTGCCCACTCAGCGAAGAGCCCAAGAATTCATCAAGGC	1298
Qy	601	TACGCTCAGAGAGCCCTCAGCGCCCGGCTAGTGGGCCAGTGCCCTTCCCAACCCCTG	660
Db	1299	TACGCTCAGAGAGCCCTCAGCGCCCGGCTAGTGGGCCAGTGCCCTTCCCAACCCCTG	1358
Qy	661	GTTCCTGAAATCCATGCTCTCGGAGGAGGCCCACTTCATCAACACCAAGCAGACTCAAGTT	720
Db	1359	GTTCCTGAAATCCATGCTCTCGGAGGAGGCCCACTTCATCAACACCAAGCAGACTCAAGTT	1418
Qy	721	CTCTCTGGCCGGAAGTGCCTGGCAGCGTTCCTGATGAGCTTACAAGTCCATCTCTGTGTAT	780
Db	1419	CTCTCTGGCCGGAAGTGCCTGGCAGCGTTCCTGATGAGCTTACAAGTCCATCTCTGTGTAT	1478
Qy	781	GGCTCCATCCCTGAGCTGTGGGGGCCCAACCGCAGCCCTTCCCTCGGCTGAGCCG	840
Db	1479	GGCTCCATCCCTGAGCTGTGGGGGCCCAACCGCAGCCCTTCCCTCGGCTGAGCCG	1538
Qy	841	GCACCTCTGCAGGAGACACTGGTGGCTCCTCCAGTTCTCTCTCTCTCTGACAACGAGCT	900
Db	1539	GCACCTCTGCAGGAGACACTGGTGGCTCCTCCAGTTCTCTCTCTCTCTGACAACGAGCT	1598
Qy	901	CGCCCCATTTGCCGGGCTAAGTCACTGCACACCTCCCGGTGACACACTCACCACTGCT	960
Db	1599	CGCCCCATTTGCCGGGCTAAGTCACTGCACACCTCCCGGTGACACACTCACCACTGCT	1658
Qy	961	GCACCCCGGGGCTTCTGTGGCCCTCGGCAGGCTGCTTGAGGAAGCCGAGGCCAGTGA	1020
Db	1659	GCACCCCGGGGCTTCTGTGGCCCTCGGCAGGCTGCTTGAGGAAGCCGAGGCCAGTGA	1718
Qy	1021	GCGCTCCACCGAGGCCACGCTCGGCTGATCTCCGAGGGTCCGGGCCACCGGCCG	1080
Db	1719	GCGCTCCACCGAGGCCACGCTCGGCTGATCTCCGAGGGTCCGGGCCACCGGCCG	1778
Qy	1081	CAAGGGTGGTGGCCCGGCACAGCTATCCGAGCTGTTCTTACCAACAGCGGGTGA	1140
Db	1779	CAAGGGTGGTGGCCCGGCACAGCTATCCGAGCTGTTCTTACCAACAGCGGGTGA	1838
Qy	1141	GAGCCCTGAGCAGGGGCCCTGGCCCGGGAGCAGCGCACCCGGCCCGGGCGGCA	1200
Db	1839	GAGCCCTGAGCAGGGGCCCTGGCCCGGGAGCAGCGCACCCGGCCCGGGCGGCA	1898
Qy	1201	CTTGCTGAAGGGCGGGTACATTGGGGGGCGCTGCGAGGCTCGCGAGCCACTGATGGA	1260
Db	1899	CTTGCTGAAGGGCGGGTACATTGGGGGGCGCTGCGAGGCTCGCGAGCCACTGATGGA	1958
Qy	1261	GCACCGGTGCTGGAGGAGGAGCGCGCCAGGAGGAGCAGGCCACCTCTGTGCCAAAGC	1320
Db	1959	GCACCGGTGCTGGAGGAGGAGCGCGCCAGGAGGAGCAGGCCACCTCTGTGCCAAAGC	2018
Qy	1321	CCCTCATTTGAGACTGCCCTCGGCTGCTGCTCTGGCACCCACTTGGCCCCCTGGCCA	1380
Db	2019	CCCTCATTTGAGACTGCCCTCGGCTGCTGCTCTGGCACCCACTTGGCCCCCTGGCCA	2078
Qy	1381	CAGCCACTCCCTGGAACATGACTCTCCGAGCACCCCGCCCTCTCTCGAGGCCCTGGC	1440
Db	2079	CAGCCACTCCCTGGAACATGACTCTCGAGCACCCCGCCCTCTCTCGAGGCCCTGGC	2138
Qy	1441	TGAGGCACAGCACTGCCCTTACGCCCTTCCGGGGGGGGCCCCCTATCAGGGACATGGGGCA	1500

Db	2139	 TGAGGCACAGCGACTGCTCTT	 CAGCCCCCTCGGGGGGGCCCTAT	 CAGGGGACATGGGGCA	2198
Qy	1501	 CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGA	1560		
Db	2199	 CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGA	2258		
Qy	1561	 GAGGCCATCCCGGACAGCCCTTTGGGGGACGACGCCCTTTCTGCGACCCCAAGCAGGG	1620		
Db	2259	 GAGGCCATCCCGGACAGCCCTTTGGGGGACGACGCCCTTTCTGCGACCCCAAGCAGGG	2318		
Qy	1621	 TTCTGCCCCCAGGAGGGCTGCAAGCCCAACCCAGCAGTTGCCCACTGCCCTCTTGCTC	1680		
Db	2319	 TTCTGCCCCCAGGAGGGCTGCAAGCCCAACCCAGCAGTTGCCCACTGCCCTCTTGCTC	2378		
Qy	1681	 CTTCCCTCAGGATCTTGCAAGAGGGCCCTTAGTACCTCAAGCCCTTTCTTGCGGACA	1740		
Db	2379	 CTTCCCTCAGGATCTTGCAAGAGGGCCCTTAGTACCTCAAGCCCTTTCTTGCGGACA	2438		
Qy	1741	 GCCCCAGGCAACCCCTGCCCTGCAAAAGCAAGCCCCCATTTGGAATCTTAAGATGGGGC	1800		
Db	2439	 GCCCCAGGCAACCCCTGCCCTGCAAAAGCAAGCCCCCATTTGGAATCTTAAGATGGGGC	2498		
Qy	1801	 TGGAGACATCTCTTCTGCGGAGGCCAAACCCCGCCCTGCGAGTTCCCAAGGCTCAGC	1860		
Db	2499	 TGGAGACATCTCTTCTGCGGAGGCCAAACCCCGCCCTGCGAGTTCCCAAGGCTCAGC	2558		
Qy	1861	 CTCCAGGCGAGCTCTTCCCAAGTGAGCTCCCTCAGGGTGGGCTCTCCAGGTTGGGCA	1920		
Db	2559	 CTCCAGGCGAGCTCTTCCCAAGTGAGCTCCCTCAGGGTGGGCTCTCCAGGTTGGGCA	2618		
Qy	1921	 AGAGCTGGGCCCTCCCTGATGCGAGGGCTGGACCCAGGAGGCTTGAGATCTGTCCGA	1980		
Db	2619	 AGAGCTGGGCCCTCCCTGATGCGAGGGCTGGACCCAGGAGGCTTGAGATCTGTCCGA	2678		
Qy	1981	 CTCCACACCCACTTGCAGCGGCTCAGGAACAGGCGACCATGCGCAAGTTCTCCCTGG	2040		
Db	2679	 CTCCACACCCACTTGCAGCGGCTCAGGAACAGGCGACCATGCGCAAGTTCTCCCTGG	2738		
Qy	2041	 TGGTCGGGGGGCTACGACGAGGTGGCTGGCTATGGCACTTTGCTTTGTTGGAGATGC	2100		
Db	2739	 TGGTCGGGGGGCTACGACGAGGTGGCTGGCTATGGCACTTTGCTTTGTTGGAGATGC	2798		
Qy	2101	 AGGGGCGATCTGGGGCAGGGGCCCATGTTGGCGCAGGATAGCCTGGGCTGTCTCCAGTC	2160		
Db	2799	 AGGGGCGATCTGGGGCAGGGGCCCATGTTGGCGCAGGATAGCCTGGGCTGTCTCCAGTC	2858		
Qy	2161	 GGAGGAGGAGGACAGGAGGAGGCCAGGCTGAGTCCAGTCGGAGGAGCAGCAGGAGGC	2220		
Db	2859	 GGAGGAGGAGGAGCAGGAGGAGGCCAGGCTGAGTCCAGTCGGAGGAGCAGCAGGAGGC	2918		
Qy	2221	 CAGGGCTGAGAGCCCACTGCCCCAGGTGAGTCGAGGAGGCTGTGCTGAGTGGCAGGGC	2280		
Db	2919	 CAGGGCTGAGAGCCCACTGCCCCAGGTGAGTCGAGGAGGCTGTGCTGAGTGGCAGGGC	2978		
Qy	2281	 TCCCAACAGGAGCTCTCCAGAGCCCAACCCCATGGAGGACATCGGGAGGCTCTCCCTGGT	2340		
Db	2979	 TCCCAACAGGAGCTCTCCAGAGCCCAACCCCATGGAGGACATCGGGAGGCTCTCCCTGGT	3038		
Qy	2341	 GCAGATCCGGGACCTGTGAGGTGATCGGAGGCGGCCGACACAAATATCCCTGGACATTT	2400		
Db	3039	 GCAGATCCGGGACCTGTGAGGTGATCGGAGGCGGCCGACACAAATATCCCTGGACATTT	3098		
Qy	2401	 CGAGGTGGACCCGGCTACTCAACTCTCAGACCTGTACGATACAAATATCCCTGGACATTT	2460		
Db	3099	 CGAGGTGGACCCGGCTACTCAACTCTCAGACCTGTACGATACAAATATCCCTGGACATTT	3158		
Qy	2461	 CGAGTTTATGATCTTCAGGAAAGTCCCAGAGTCCGCTCAGGAGGCGCCCTCCCCAT	2520		
Db	3159	 CGAGTTTATGATCTTCAGGAAAGTCCCAGAGTCCGCTCAGGAGGCGCCCTCCCCAT	3218		
Qy	2521	 GGCTGAGGAGGAGCTGGCGAGTTTCCGAGGCCCACTGGGCCCTGGCGAGGTGAATCGG	2580		


```
Db 3219 GGCTGAGGAGGAGCTGGCCGAGTTCCCGAGGCCACCTGGCCCTGGCCAGGTGAACCTGGG 3278
Qy 2581 CCCCCACGCGAGCTTGAGATACAGAGGAGTACAGAGGATGAGACGCGTGTCTGGCAGA 2640
Db 3279 CCCCCACGCGAGCTTGAGATACAGAGGAGTACAGAGGATGAGACGCGTGTCTGGCAGA 3338
Qy 2641 GGCTGCGGTGGCAGGAGCGCAAGTGGTCTCGCCGTCACGACGCTCTTCCACTTCCC 2700
Db 3339 GGCTGCGGTGGCAGGAGCGCAAGTGGTCTCGCCGTCACGACGCTCTTCCACTTCCC 3398
Qy 2701 TGGGAGGACACCTGCGCGTGGATGAGCCCTGCAGAGCTGGGGTGGCTGAGAGAGTGAAGGC 2760
Db 3399 TGGGAGGACACCTGCGCGTGGATGAGCCCTGCAGAGCTGGGGTGGCTGAGAGAGTGAAGGC 3458
Qy 2761 CTCCGTGAGACATCTCCCGGATCTGAAGGAGCGCCGAGAGTCTGGAGAGAGAGG 2820
Db 3459 CTCCGTGAGACATCTCCCGGATCTGAAGGAGCGCCGAGAGTCTGGAGAGAGAGG 3518
Qy 2821 GCCCCCAGGAGAGAGCCAGCCCTTGTCTTCCCTCCGCGCTCAGGTCTGAAGAGCTGGA 2880
Db 3519 GCCCCCAGGAGAGAGCCAGCCCTTGTCTTCCCTCCGCGCTCAGGTCTGAAGAGCTGGA 3578
Qy 2881 CCGAGCGCCGACATCTCTAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCAGT 2940
Db 3579 CCGAGCGCCGACATCTCTAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCAGT 3638
Qy 2941 GACACTGCGCTGCGAGGTGTGAGCCGAGCTGCGCCAGCCACCTGGAGCAAGAGCG 3000
Db 3639 GACACTGCGCTGCGAGGTGTGAGCCGAGCTGCGCCAGCCACCTGGAGCAAGAGCG 3698
Qy 3001 A 3001
Db 3699 A 3699

RESULT 2
US-10-274-978-1
; Sequence 1, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CLO00927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Human
US-10-274-978-1

Query Match 99.9%; Score 2997.8; DB 4; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTACGAGCAGAACTCGGGCCAGGATACAGAGGAGCGAGACATCTTGGCGCGGTGAG 60
Db 699 CCTACGAGCAGAACTCGGGCCAGGATACAGAGGAGCGAGACATCTTGGCGCGGTGAG 758
Qy 61 CCACCCGCTGTACAGGGGCTGTGGACCACTTTGAGACCCGCAAGACCTCATCTCTCAT 120
Db 759 CCACCCGCTGTACAGGGGCTGTGGACCACTTTGAGACCCGCAAGACCTCATCTCTCAT 818
Qy 121 CCTGAGCTGTGTATCCGAGGAGCTGTGGACCGCTGTACAGGAGGCGGTGTGAC 180
```

```
Db 819 CCTGAGAGCTGTCTCATCTCCGAGGAGCTGTGGACCGCCTGTACAGGAAGGCGGTGCTGAC 878
Qy 181 GGAGGCCGAGGTCAAGGTCTACATCCAGAGAGCTGGTGGAGGGGTGCACTACTTGCACAG 240
Db 879 GGAGGCCGAGGTCAAGGTCTACATCCAGAGAGCTGGTGGAGGGGTGCACTACTTGCACAG 938
Qy 241 CCATGGGGTCTCCACCTGGACATAAAGCCCTCTAAACATCCTGATGGTGCATCTGCCCG 300
Db 939 CCATGGGGTCTCCACCTGGACATAAAGCCCTCTAAACATCCTGATGGTGCATCTGCCCG 998
Qy 301 GGAAGACATTAATAATCTGCCACTTTTGGCTTTTGGCCAGAACATCAACCCAGCAGAGCTGCA 360
Db 999 GGAAGACATTAATAATCTGCCACTTTTGGCTTTTGGCCAGAACATCAACCCAGCAGAGCTGCA 1058
Qy 361 GTTCAGCAGTACCGCTCCCTGAGTTCTGCTCCCGGAGATCATTCAGCAGAAACCTGT 420
Db 1059 GTTCAGCAGTACCGCTCCCTGAGTTCTGCTCCCGGAGATCATTCAGCAGAAACCTGT 1118
Qy 421 GAGCGAGGCTCCGACATTTGGGCCATGGGTGTCTCTTACCTCAGCTCAGCTGCTC 480
Db 1119 GAGCGAGGCTCCGACATTTGGGCCATGGGTGTCTCTTACCTCAGCTCAGCTGCTC 1178
Qy 481 ATCCCCATTTCCCGCGCAGAGTGAACCGTGGCCACCTCTCTGAACCTCTTGGAGGGCGCGT 540
Db 1179 ATCCCCATTTCCCGCGCAGAGTGAACCGTGGCCACCTCTCTGAACCTCTTGGAGGGCGCGT 1238
Qy 541 GTCATGAGCAGGCCCCATGCTGCCACCTCAGCGAAGACGCCAAAGACTTTCATCAAGGC 600
Db 1239 GTCATGAGCAGGCCCCATGCTGCCACCTCAGCGAAGACGCCAAAGACTTTCATCAAGGC 1298
Qy 601 TAGCTCAGAGAGGCCCCCTCAGGCCCGGCTAGTGGCGGCCAGTGCCTCTCCACCCCTG 660
Db 1299 TAGCTCAGAGAGGCCCCCTCAGGCCCGGCTAGTGGCGGCCAGTGCCTCTCCACCCCTG 1358
Qy 661 GTTCCTGAATCCATGCTCGGAGGAGGCGCCACTTCATCAACCAAGCAGCTCAAGTT 720
Db 1359 GTTCCTGAATCCATGCTCGGAGGAGGCGCCACTTCATCAACCAAGCAGCTCAAGTT 1418
Qy 721 CTCTCTGCGCCGAAAGTCTGCTGGCAGCTTCCCTGATGAGCTTCAAGTTCATCTCTGTTAT 780
Db 1419 CTCTCTGCGCCGAAAGTCTGCTGGCAGCTTCCCTGATGAGCTTCAAGTTCATCTCTGTTAT 1478
Qy 781 GCGTCTCATCTCTGAGCTGCTGGGGGCCACCCGACAGAGCCCTCCCTCGCGGTAGCCCG 840
Db 1479 GCGTCTCATCTCTGAGCTGCTGGGGGCCACCCGACAGAGCCCTCCCTCGCGGTAGCCCG 1538
Qy 841 GCACCTCTGAGGAGCACTGTGTGCTCTCTCAGTTCCTCTCTCTCTGACAAAGAGCT 900
Db 1539 GCACCTCTGAGGAGCACTGTGTGCTCTCTCAGTTCCTCTCTCTCTGACAAAGAGCT 1598
Qy 901 CGCCCCATTTGCGCGGCTAAGTCACTGACACCTCTCCCGGTGACACTCACCACCTGCT 960
Db 1599 CGCCCCATTTGCGCGGCTAAGTCACTGACACCTCTCCCGGTGACACTCACCACCTGCT 1658
Qy 961 GCACCCCGGGGCTTCTCGGGCCCTCGGCGAGCTGCTGAGGAAGCCGAGGCGCAGTGA 1020
Db 1659 GCACCCCGGGGCTTCTCGGGCCCTCGGCGAGCTGCTGAGGAAGCCGAGGCGCAGTGA 1718
Qy 1021 GCGTCTCAACGAGGCGCCAGCTCCGCTGCTATCTCCGAGGAGTCCGCGGCCACCGCGCGC 1080
Db 1719 GCGTCTCAACGAGGCGCCAGCTCCGCTGCTATCTCCGAGGAGTCCGCGGCCACCGCGCGC 1778
Qy 1081 CCAGGGTGTGCTGCCCGGACAGCTCATCCGAGCTCTTCTTCAACAGGCGGGTGA 1140
Db 1779 CCAGGGTGTGCTGCCCGGACAGCTCATCCGAGCTCTTCTTCAACAGGCGGGTGA 1838
Qy 1141 GAGCCCTGAGCAGGGGCGCTGGCCCGGGGAGCAGCGGCGACCCGCGCGCGCGCGCA 1200
Db 1839 GAGCCCTGAGCAGGGGCGCTGGCCCGGGGAGCAGCGGCGACCCGCGCGCGCGCGCA 1898
Qy 1201 CTTCTGAAGGGGGGTACATTGGGGGGGCGCTCCAGGCTGCGCGAGGCGCTGATGGA 1260
Db 1899 CTTCTGAAGGGGGGTACATTGGGGGGGCGCTCCAGGCTGCGCGAGGCGCTGATGGA 1958
```


Query Match				
Beat Local Similarity 99.9%; Score 2997.8; DB 4; Length 5207;				
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;				
Qy	1	CCTACGGAGCAGAACTCGGGCCCGAGGCAACAGGAGGAGAGACATCTCTGGCGCGCTGAG	60	
Db	699	CCTACGGAGCAGAACTCGGGCCCGAGGCAACAGGAGGAGAGACATCTCTGGCGCGCTGAG	758	
Qy	61	CCACCCGCTGTACGGGGCTGTGGACAGTTTGTAGACCCGCAAGACCTTCATCTCTCAT	120	
Db	759	CCACCCGCTGTGTACGGGGCTGTGGACAGTTTGTAGACCCGCAAGACCTTCATCTCTCAT	818	
Qy	121	CCTGAGCTGTCTATCCGAGGAGCTGTGACCGCTGTACAGGAAGGCGGTGTGAC	180	
Db	819	CCTGAGCTGTCTATCCGAGGAGCTGTGACCGCTGTACAGGAAGGCGGTGTGAC	878	
Qy	181	GGAGCCGAGGTCAAGGTCTACATCCAGACGCTGTGTGAGGGGCTGCACTACCTGCACAG	240	
Db	879	GGAGCCGAGGTCAAGGTCTACATCCAGACGCTGTGTGAGGGGCTGCACTACCTGCACAG	938	
Qy	241	CCATGGGTTTCCACCTGTGACATAAAGCCCTCTAAACATCTGTATGTTGTGATCTGTGCCG	300	
Db	939	CCATGGGTTTCCACCTGTGACATAAAGCCCTCTAAACATCTGTATGTTGTGATCTGTGCCG	998	
Qy	301	GGAGACATTAATAATCTGCGACTTTGGCTTTGCCAGAACATCAACCCAGCAGAGCTGCA	360	
Db	999	GGAGACATTAATAATCTGCGACTTTGGCTTTGCCAGAACATCAACCCAGCAGAGCTGCA	1058	
Qy	361	GTTCAGCCAGTACCGCTCCCTGAGTTCTGTCTCCCGGAGATCATCCAGCAGAAACCTGT	420	
Db	1059	GTTCAGCCAGTACCGCTCCCTGAGTTCTGTCTCCCGGAGATCATCCAGCAGAAACCTGT	1118	
Qy	421	GAGGAAGCCTCCGACATTTGGGCAATGGGTGTCTATCTCTACCTCAGCTGACCTGTCTC	480	
Db	1119	GAGGAAGCCTCCGACATTTGGGCAATGGGTGTCTATCTCTACCTCAGCTGACCTGTCTC	1178	
Qy	481	ATCCCCATTTGCCGCGAGAGTACCGTGCACACCTCTCTGAACCTCTCTGAGGGGCGGT	540	
Db	1179	ATCCCCATTTGCCGCGAGAGTACCGTGCACACCTCTCTGAACCTCTCTGAGGGGCGGT	1238	
Qy	541	GTCTAGAGCAGCCCATGGCTGGCCACCTCAGCGAAGACGCCAAAGACTTTCATCAAGGC	600	
Db	1239	GTCTAGAGCAGCCCATGGCTGGCCACCTCAGCGAAGACGCCAAAGACTTTCATCAAGGC	1298	
Qy	601	TACGCTGAGAGAGCCCTCAGGCGCGGCTAGTGTGCGGCCAGTGTCTCTCCCAACCCCTG	660	
Db	1299	TACGCTGAGAGAGCCCTCAGGCGCGGCTAGTGTGCGGCCAGTGTCTCTCCCAACCCCTG	1358	
Qy	661	GTCTCTGAATCCATGCTCGGAGGAGGCCCACTTCATCAACCAAGCAGCTCAAGTT	720	
Db	1359	GTCTCTGAATCCATGCTCGGAGGAGGCCCACTTCATCAACCAAGCAGCTCAAGTT	1418	
Qy	721	CCTCTGCGCCGAAAGTCTGCGCAGCGTTCCCTGATGAGCTACAAAGTCCATCTCTGGTGAT	780	
Db	1419	CCTCTGCGCCGAAAGTCTGCGCAGCGTTCCCTGATGAGCTACAAAGTCCATCTCTGGTGAT	1478	
Qy	781	CGCTTCCATCCTGAGTGTCTGCGGGCCCAACCCGACAGCCCTCTCCCTCGCGGTAGCCCG	840	
Db	1479	CGCTTCCATCCTGAGTGTCTGCGGGCCCAACCCGACAGCCCTCTCCCTCGCGGTAGCCCG	1538	
Qy	841	GCACCTCTGACGGACACTGTGTGCTCTCCAGTTCTCTCTCTCTCTGACACAGAGCT	900	
Db	1539	GCACCTCTGACGGACACTGTGTGCTCTCCAGTTCTCTCTCTCTCTCTGACACAGAGCT	1598	
Qy	901	CGCCCCATTTGCCCGGGCTAAGTCACTGCCACCTCTCCCGGTGTACACTCACCACTGTCT	960	
Db	1599	CGCCCCATTTGCCCGGGCTAAGTCACTGCCACCTCTCCCGGTGTACACTCACCACTGTCT	1658	
Qy	961	GCACCCCCGGGCTTCTCGGCGCTCTGGCCAGCTGTCTGAGGAAGCCGAGGCCAGTGA	1020	
Db	1659	GCACCCCCGGGCTTCTCGGCGCTCTGGCCAGCTGTCTGAGGAAGCCGAGGCCAGTGA	1718	
Qy	1021	CGGCTCCACCGAGGCCCTCCGCTGTCTCCCGAGGGGTGCGGGGCCACCGGGCGC	1080	

Db	1719	GGGCTCCACCGAGGCCCCAGCTCCGCTCTGCATCTCCCGAGGGTCCCGGGCCACCGCGCGC	1778	
Qy	1081	CCAGGGCTGCGTGCCTCCCGGCACACGCGTCACTCGCAGCCTGTTCTTACCAACAGGGGGTGA	1140	
Db	1779	CCAGGGCTGCGTGCCTCCCGGCACACGCGTCACTCGCAGCCTGTTCTTACCAACAGGGGGTGA	1838	
Qy	1141	GAGCCCTGAGCAGCGGGCCCTTGCCCCCGGGGAGCAGGGGCAACCGGGCCCGGGCGGCA	1200	
Db	1839	GAGCCCTGAGCAGCGGGCCCTTGCCCCCGGGGAGCAGGGGCAACCGGGCCCGGGCGGCA	1898	
Qy	1201	CCTCTGAAGGGGGGTACATTGGGGGGCGCTGCCAGGCGCTGCGCAGCCACTGTATGGA	1260	
Db	1899	CCTCTGAAGGGGGGTACATTGGGGGGCGCTGCCAGGCGCTGCGCAGCCACTGTATGGA	1958	
Qy	1261	GCACCGGCTGTGTGAGGAGGCGGCCAGGGAGGAGCAGGCCACCTCTCTGGGCCAAAGC	1320	
Db	1959	GCACCGGCTGTGTGAGGAGGCGGCCAGGGAGGAGCAGGCCACCTCTCTGGGCCAAAGC	2018	
Qy	1321	CCCTCATTTGAGACTGCGCTCCGCGTGCCTCTCTGGCACCACCTTGGGCCCTTGCGCA	1380	
Db	2019	CCCTCATTTGAGACTGCGCTCCGCGTGCCTCTCTGGCACCACCTTGGGCCCTTGCGCA	2078	
Qy	1381	CAGCACTCTCCTGAACATGACTCTCCGAGCACCCCGCCCTCTCTCGAGGGCTTGCGG	1440	
Db	2079	CAGCACTCTCCTGAACATGACTCTCCGAGCACCCCGCCCTCTCTCGAGGGCTTGCGG	2138	
Qy	1441	TGAGGCACAGCGACTGCTTTCAGCCCCCTCGGGGGGGGCCCTTATCAGGGACATGGGGCA	1500	
Db	2139	TGAGGCACAGCGACTGCTTTCAGCCCCCTCGGGGGGGGCCCTTATCAGGGACATGGGGCA	2198	
Qy	1501	CCCTCAGGGTCCAAAGCAGCTTCCATCCTCTGTGGGCAACCCAGGCACTGTCTAGCCAGA	1560	
Db	2199	CCCTCAGGGTCCAAAGCAGCTTCCATCCTCTGTGGGCAACCCAGGCACTGTCTAGCCAGA	2258	
Qy	1561	GAGGCATCCCGGACAGCCCTTGGGGGAGCAGCCAGCCCTTCTGCGCACCCAAAGCAGGG	1620	
Db	2259	GAGGCATCCCGGACAGCCCTTGGGGGAGCAGCCAGCCCTTCTGCGCACCCAAAGCAGGG	2318	
Qy	1621	TTCTGCCCCCGAGGAGGCTGACGCCCCCAGCCAGCAGTTGCCCCATGCCCCCTCTGGGCTC	1680	
Db	2319	TTCTGCCCCCGAGGAGGCTGACGCCCCCAGCCAGCAGTTGCCCCATGCCCCCTCTGGGCTC	2378	
Qy	1681	CTTCCCTCCAGGATCTTTGCAAGAGGCCCCCTTAGTACCTCTCAAGCCCTTCTTGGGCA	1740	
Db	2379	CTTCCCTCCAGGATCTTTGCAAGAGGCCCCCTTAGTACCTCTCAAGCCCTTCTTGGGCA	2438	
Qy	1741	GCCCCAGGCACCCCTCCCTGCGCAAGCAAGCCCCCCTTGGACTCTAAGATTGGGGCC	1800	
Db	2439	GCCCCAGGCACCCCTCCCTGCGCAAGCAAGCCCCCCTTGGACTCTAAGATTGGGGCC	2498	
Qy	1801	TGGAGACATCTCTTCTTGGAGGCCAAAACCCGGGCCCTTGACAGTTCCCCAGGGTTCAGC	1860	
Db	2499	TGGAGACATCTCTTCTTGGAGGCCAAAACCCGGGCCCTTGACAGTTCCCCAGGGTTCAGC	2558	
Qy	1861	CTCCCAAGGAGCTCTTCCCAAGTACCTCCTCAGGGTGGGCTCTCTCCAGGTGGGCAC	1920	
Db	2559	CTCCCAAGGAGCTCTTCCCAAGTACCTCCTCAGGGTGGGCTCTCTCCAGGTGGGCAC	2618	
Qy	1921	AGAGCCTGGCCCCCTCCCTGGAATGCGAGGGCTGGACCCAGAGGCTGAGGATCTGTCGA	1980	
Db	2619	AGAGCCTGGCCCCCTCCCTGGAATGCGAGGGCTGGACCCAGAGGCTGAGGATCTGTCGA	2678	
Qy	1981	CTCCACACCCACCTTTCAGCGGGCTCAGGAACAGGCGACCATGCGCAAGTTTCTCCCTGGG	2040	
Db	2679	CTCCACACCCACCTTTCAGCGGGCTCAGGAACAGGCGACCATGCGCAAGTTTCTCCCTGGG	2738	
Qy	2041	TGCTCGGGGGGTACGACAGGCGTGTGGCTATGCGCACTTTGCTTTTGGTGGAGATGC	2100	
Db	2739	TGCTCGGGGGGTACGACAGGCGTGTGGCTATGCGCACTTTGCTTTTGGTGGAGATGC	2798	
Qy	2101	AGGGGGCATCTGGGGGAGGGGCCCATGTGGGCGAGGATAGCCTGGGCTGTGTCCTAGTC	2160	


```
Db 2799 AGGGGGCATGCTGGGGCAGGGGCCCATGTGTGGCCAGGATAGCCTGGGCTGTGTCCAGTC 2858
Qy 2161 GGAGGAGGAGCAGCAGAGGAGGGCAGGGCTGAGTCCAGTTCGAGAGCAGCAGGAGGC 2220
Db 2859 GGAGGAGGAGCAGCAGAGGAGGGCAGGGCTGAGTCCAGTTCGAGAGCAGCAGGAGGC 2318
Qy 2221 CAGGGCTGAGAGGCCACTGCCCCCAGGTCAGTGCAAGGCCCTGTGCTGAGGTTCGGCAGGGC 2280
Db 2919 CAGGGCTGAGAGGCCACTGCCCCCAGGTCAGTGCAAGGCCCTGTGCTGAGGTTCGGCAGGGC 2978
Qy 2281 TCCACACAGAGCTCTCCAGAGCCCAACCCATGGGAGGACATCGGGCAGGTCTCCCTGGT 2340
Db 2979 TCCACACAGAGCTCTCCAGAGCCCAACCCATGGGAGGACATCGGGCAGGTCTCCCTGGT 3038
Qy 2341 GCAGATCCGGACCTGTCAAGTATGCGAGGGGGCCGACACAAATATCCCTGGACATTTTC 2400
Db 3039 GCAGATCCGGACCTGTCAAGTATGCGAGGGGGCCGACACAAATATCCCTGGACATTTTC 3098
Qy 2401 CGAGGTGACCCCGCCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
Db 3099 CGAGGTGACCCCGCCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 3158
Qy 2461 CGAGTTTATGATCTTACGAAAGTCCCAAGTCCGCTCAGCAGAGCCGCCCTCCCCCAT 2520
Db 3159 CGAGTTTATGATCTTACGAAAGTCCCAAGTCCGCTCAGCAGAGCCGCCCTCCCCCAT 3218
Qy 2521 GGCTGAGGAGAGCTGGCCGAGTTCCCGGAGCCCAAGTCCGCTCAGCAGAGCCGCCCTCCCCCAT 2580
Db 3219 GGCTGAGGAGAGCTGGCCGAGTTCCCGGAGCCCAAGTCCGCTCAGCAGAGCCGCCCTCCCCCAT 3278
Qy 2581 CCCCACGCGAGGCTGGAGATCAAGAGGAGTCAAGAGGATGTGGACCGCTGTCTGGCAGA 2640
Db 3279 CCCCACGCGAGGCTGGAGATCAAGAGGAGTCAAGAGGATGTGGACCGCTGTCTGGCAGA 3338
Qy 2641 GGCTGCGTGGGAGGAGGAGGAGGAGTGTCTGCGCTCAGCAGGCTCTTCCACTTCCC 2700
Db 3339 GGCTGCGTGGGAGGAGGAGGAGGAGTGTCTGCGCTCAGCAGGCTCTTCCACTTCCC 3398
Qy 2701 TGGAGAGCAGCTGCGCTGGATGAGCTGCAGAGCTGGGCTGGTGCAGAGAGTGAAGGC 2760
Db 3399 TGGAGAGCAGCTGCGCTGGATGAGCTGCAGAGCTGGGCTGGTGCAGAGAGTGAAGGC 3458
Qy 2761 CTCGTTGGAGCAGATCTCCCGGATCTCTAAGGGCAGGCCCGGAAGGTTCGAGAGAGGAGG 2820
Db 3459 CTCGTTGGAGCAGATCTCCCGGATCTCTAAGGGCAGGCCCGGAAGGTTCGAGAGAGGAGG 3518
Qy 2821 GCGCCCGCAGGAGAGCAGGCTTGTCTCTTCCGCTCTCAGGTCTGAGAGCTGGGA 2880
Db 3519 GCGCCCGCAGGAGAGCAGGCTTGTCTCTTCCGCTCTCAGGTCTGAGAGCTGGGA 3578
Qy 2881 CCGAGCGCCGACATTTCTAAGGGAGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 2940
Db 3579 CCGAGCGCCGACATTTCTAAGGGAGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 3638
Qy 2941 GACATCGCTTCGAGGTGTACGCCAGCAGCTGCCAGGCCACCTTGGAGCAAGACGG 3000
Db 3639 GACATCGCTTCGAGGTGTACGCCAGCAGCTGCCAGGCCACCTTGGAGCAAGACGG 3698
Qy 3001 A 3001
Db 3699 A 3699
```

RESULT 4
US-10-697-263-1
; Sequence 1, Application US/10697263
; Patent No. 6812014
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV2

```
; CURRENT APPLICATION NUMBER: US/10/697,263  
; CURRENT FILING DATE: 2003-10-31  
; PRIOR APPLICATION NUMBER: 10/274,978  
; PRIOR FILING DATE: 2002-10-22  
; PRIOR APPLICATION NUMBER: 09/858,664  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 09/711,134  
; PRIOR FILING DATE: 2000-11-14  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 5207  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-697-263-1  
  
Query Match 99.9%; Score 2997.8; DB 4; Length 5207;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
Qy 1 CCTACCGAGCAGAACTTCGGGCCCCAGGCATACAGGGAGCGAGACATCTCTGGCCGCGTGGAG 60  
Db 699 CCTACCGAGCAGAACTTCGGGCCCCAGGCATACAGGGAGCGAGACATCTCTGGCCGCGTGGAG 758  
Qy 61 CCAACCCGCTGGTCAACGGGGCTGTGGACCAAGTTTGAGACCCGCAAGACCCCTCATCTCAT 120  
Db 759 CCAACCCGCTGGTCAACGGGGCTGTGGACCAAGTTTGAGACCCGCAAGACCCCTCATCTCAT 818  
Qy 121 CTGTGAGCTGTGCTCATCCGAGGAGCTGTGGACCCGCTGTACAGGAAGGGCGTGGTGAC 180  
Db 819 CTGTGAGCTGTGCTCATCCGAGGAGCTGTGGACCCGCTGTACAGGAAGGGCGTGGTGAC 878  
Qy 181 GGAGGCCGAGGTCAGAGGCTTACATCCAGCAGCTGTGTGGAGGGGCTGCACCTACCTGCACAG 240  
Db 879 GGAGGCCGAGGTCAGAGGCTTACATCCAGCAGCTGTGTGGAGGGGCTGCACCTACCTGCACAG 938  
Qy 241 CCATGCGCTTCCACCTGAGCATAAAGCCCTCTAAACATCTGATGGTGGATCTCTGCCCCG 300  
Db 939 CCATGCGCTTCCACCTGAGCATAAAGCCCTCTAAACATCTGATGGTGGATCTCTGCCCCG 998  
Qy 301 GGAAGACATTAATAATCTGGGACCTTTCGCTTTGGCCAGAGACATCACCCAGCAGAGCTGCA 360  
Db 999 GGAAGACATTAATAATCTGGGACCTTTCGCTTTGGCCAGAGACATCACCCAGCAGAGCTGCA 1058  
Qy 361 GTTCAGCCAGTACGGCTCCCTGAGTTCCCTCCCGAGATCATCCAGCAGAACCTCTGT 420  
Db 1059 GTTCAGCCAGTACGGCTCCCTGAGTTCCCTCCCGAGATCATCCAGCAGAACCTCTGT 1118  
Qy 421 GAGCGAAGCCTCCGACATTTGGGCCATGGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 480  
Db 1119 GAGCGAAGCCTCCGACATTTGGGCCATGGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTC 1178  
Qy 481 ATCCCATTTTCGGGGAGAGTGACCGTGCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 540  
Db 1179 ATCCCATTTTCGGGGAGAGTGACCGTGCACCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 1238  
Qy 541 GTCATGAGCAGCCCATGCTGCCACCTCAGCGAAGACGCCCAAGACTTCTCAAGGC 600  
Db 1239 GTCATGAGCAGCCCATGCTGCCACCTCAGCGAAGACGCCCAAGACTTCTCAAGGC 1298  
Qy 601 TAGCTGCGAGAGCCCTCTCAGGCCCGGCTAGTGGGGCCAGTCTCTCTCTCTCTCTCTCTCTCT 660  
Db 1299 TAGCTGCGAGAGCCCTCTCAGGCCCGGCTAGTGGGGCCAGTCTCTCTCTCTCTCTCTCTCT 1358  
Qy 661 GTTCTCTGAATCCATGCTGCGGAGGAGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 720  
Db 1359 GTTCTCTGAATCCATGCTGCGGAGGAGGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1418  
Qy 721 CCTCTGCGCCGAGGTCGCTGGCAGCGGTTCCCTGTAGTACGCTACAAGTCTCACTCTCTCTCT 780  
Db 1419 CCTCTGCGCCGAGGTCGCTGGCAGCGGTTCCCTGTAGTACGCTACAAGTCTCACTCTCTCT 1478  
Qy 781 GCGCTCATCCCTGAGTGTCTGGGGGCCCAACCCGACAGCCCTCTCTCTCTCTCTCTCTCTCT 840
```


|||||
1479 GGGCTCCATCCCTGAGCTGCTGGGGGGCCACCGACAGCCCTCCCTCGGGGTAGCCCG 1538
QY
841 GCACTCTGCAAGGACAATGTTGGTCTCTCAGATTCCTCTCTCTCTGACAAAGAGCT 900
Db
1539 GCACCTCTGCAAGGACAATGTTGGTCTCTCAGATTCCTCTCTCTCTGACAAAGAGCT 1598
QY
901 CGCCCCATTTGCCCGGGCTAAGTCACTGCGCAACCTCCCGGTGACACACTCACCCTGCT 960
Db
1599 GCCCCATTTGCCCGGGCTAAGTCACTGCGCAACCTCCCGGTGACACACTCACCCTGCT 1658
QY
961 GCACCCCGGGGCTTTCTGCGGCTCCCGGCGAGCTGCTGAGGAAGCGGAGGCGAGTGA 1020
Db
1659 GCACCCCGGGGCTTTCTGCGGCTCCCGGCGAGCTGCTGAGGAAGCGGAGGCGAGTGA 1718
QY
1021 GGGTCTCAACGAGGCCAGCTCGGCTCGCTGCAATCTCCGAGGGTGCAGGCGCGC 1080
Db
1719 GGGTCTCAACGAGGCCAGCTCGGCTCGCTGCAATCTCCGAGGGTGCAGGCGCGC 1778
QY
1081 CCAGGGCTGCTGCCCGGCGACAGCTCAATCCGACAGCTGTTTACACACAGGGCGGGTGA 1140
Db
1779 CCAGGGCTGCTGCCCGGCGACAGCTCAATCCGACAGCTGTTTACACACAGGGCGGGTGA 1838
QY
1141 GAGCCCTGAGCAGCGGGGCTTGGGCTCCCGGGGAGCAGCGGCGACCCGCGCCCGGCGCGCA 1200
Db
1839 GAGCCCTGAGCAGCGGGGCTTGGGCTCCCGGGGAGCAGCGGCGACCCGCGCCCGGCGCGCA 1898
QY
1201 CTTGCTGAAGGGCGGGTACATTTGGGGGGCGCTGCCAGGCTGCGCGAGCCACTGATGGA 1260
Db
1899 CTTGCTGAAGGGCGGGTACATTTGGGGGGCGCTGCCAGGCTGCGCGAGCCACTGATGGA 1958
QY
1261 GCACGGCTGCTGGAGGAGGGCGCCAGGGAGGAGCAGCGCCACCTTGGGCCCAAGC 1320
Db
1959 GCACGGCTGCTGGAGGAGGGCGCCAGGGAGGAGCAGCGCCACCTTGGGCCCAAGC 2018
QY
1321 CCCCTCATTTGAGACTGCGCTCCGGCTGCTGCGCTGCGCAACCACTTGGGCCCTGGCA 1380
Db
2019 CCCCTCATTTGAGACTGCGCTCCGGCTGCTGCGCTGCGCAACCACTTGGGCCCTGGCA 2078
QY
1381 CAGCCACTCTCTGAAACATGACTCTCCGAGCAACCCCGGCTCTCTGAGGGCTGCGG 1440
Db
2079 CAGCCACTCTCTGAAACATGACTCTCCGAGCAACCCCGGCTCTCTGAGGGCTGCGG 2138
QY
1441 TGAGGCACAGCACTGCTTCAAGCCCTCCGGGGGGGGCCCTATCAGGGAATGGGGCA 1500
Db
2139 TGAGGCACAGCACTGCTTCAAGCCCTCCGGGGGGGGCCCTATCAGGGAATGGGGCA 2198
QY
1501 CCCTCAGGGCTCCAGAGCACTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGA 1560
Db
2199 CCCTCAGGGCTCCAGAGCACTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGA 2258
QY
1561 GAGGCCATTCCTCCGACAGCCCTTGGGGGAGCCAGCCCTTTTCTGCCACCCCAAGCAGGG 1620
Db
2259 GAGGCCATTCCTCCGACAGCCCTTGGGGGAGCCAGCCCTTTTCTGCCACCCCAAGCAGGG 2318
QY
1621 TTCTGCCCCCAGAGGGCTCAGCCCCCAGCCAGAGTTCGCCAGCTCCCTCTGGCTC 1680
Db
2319 TTCTGCCCCCAGAGGGCTCAGCCCCCAGCCAGAGTTCGCCAGCTCCCTCTGGCTC 2378
QY
1681 CTTCCCTCCAGGATCTTGAAGAGGGCCCTTAGTACCTTCAAGCCCTTCTTGGGACA 1740
Db
2379 CTTCCCTCCAGGATCTTGAAGAGGGCCCTTAGTACCTTCAAGCCCTTCTTGGGACA 2438
QY
1741 GCCCCAGGCACCCCTCTGCTGCAAGAGCCGCCCTTATGGAATCTAAGATGGGGCC 1800
Db
2439 GCCCCAGGCACCCCTCTGCTGCAAGAGCCGCCCTTATGGAATCTAAGATGGGGCC 2498
QY
1801 TGGAGACATCTCTTCTCTGGAGGCGCAAAACCCGGGCCCTTGCAGTTCCCGAGGGTCAGC 1860
Db
2499 TGGAGACATCTCTTCTCTGGAGGCGCAAAACCCGGGCCCTTGCAGTTCCCGAGGGTCAGC 2558
QY
1861 CTCCAGGCGAGCTCTTCCCAAGTGAAGCTCCCTCAGGGTGGGCTCCCTCCAGGTGGGCAC 1920
|||||

Db
2559 CTCCAGGCGAGCTCTTCCCAAGTGAAGTCCCTCAGGGTGGGCTCCTCCAGGTGGGCAC 2618
QY
1921 AGAGCTTGCGCCCTCCCTCTGGATGCGAGGGCTGAGCCAGAGGCTGAGGATCTGTCCGA 1980
Db
2619 AGAGCTTGCGCCCTCCCTCTGGATGCGAGGGCTGAGCCAGAGGCTGAGGATCTGTCCGA 2678
QY
1981 CTCCACACCCACCTTTCAGAGCGGCTCAGGACACGCGACCATGCGCAAGTTCCTCCCTGGG 2040
Db
2679 CTCCACACCCACCTTTCAGAGCGGCTCAGGACACGCGACCATGCGCAAGTTCCTCCCTGGG 2738
QY
2041 TGGTTCGCGGGGGCTTACGACAGGCTGCGCTGCGCTATGCGACCTTTTCCCTTGGTGGAGATGC 2100
Db
2739 TGGTTCGCGGGGGCTTACGACAGGCTGCGCTGCGCTATGCGACCTTTTCCCTTGGTGGAGATGC 2798
QY
2101 AGGGGGCATGCTGGGGGACGGGGCCCATGTTGGGCCAGAGATAGCTTGGGTGTGCCAGATC 2160
Db
2799 AGGGGGCATGCTGGGGGACGGGGCCCATGTTGGGCCAGAGATAGCTTGGGTGTGCCAGATC 2858
QY
2161 GGAGGAGGAGGAGCAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2220
Db
2859 GGAGGAGGAGGAGCAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2918
QY
2221 CAGGGCTGAGAGGCCACTTGGCCCCCAGGTCAGTGCAGAGGCTGTGTGCTGAGGTTCGGCAGGGC 2280
Db
2919 CAGGGCTGAGAGGCCACTTGGCCCCCAGGTCAGTGCAGAGGCTGTGTGCTGAGGTTCGGCAGGGC 2978
QY
2281 TCCCAACAGAGAGCTCTCCAGAGCCCAACCCATGAGGAGACATCGGGCAGGTCTCCCTGGT 2340
Db
2979 TCCCAACAGAGAGCTCTCCAGAGCCCAACCCATGAGGAGACATCGGGCAGGTCTCCCTGGT 3038
QY
2341 GCAGATCCGGGACCTGTGAGTGTGCGGAGGGCGCGACACATATATCCCTGGACATTTTC 2400
Db
3039 GCAGATCCGGGACCTGTGAGTGTGCGGAGGGCGCGACACATATATCCCTGGACATTTTC 3098
QY
2401 CGAGGTGGAACCCCGCTTACCTCAACCTCTCAGACCTGTAGCATATCAAGTACCTCCCATTT 2460
Db
3099 CGAGGTGGAACCCCGCTTACCTCAACCTCTCAGACCTGTAGCATATCAAGTACCTCCCATTT 3158
QY
2461 CGAGTTTATGATTTTCAAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCGGCTCCCTCCCAT 2520
Db
3159 CGAGTTTATGATTTTCAAGGAAAGTCCCAAGTCCGCTCAGCCAGAGCGGCTCCCTCCCAT 3218
QY
2521 GGCTGAGGAGGAGTGGCGGAGTTCGGGAGCCCACTGCGCCCTGCGCCAGGTGAACCTGGG 2580
Db
3219 GGCTGAGGAGGAGTGGCGGAGTTCGGGAGCCCACTGCGCCCTGCGCCAGGTGAACCTGGG 3278
QY
2581 CCCCCAGCAGGCTGAGAGATCAGAGAGGAGTCAAGAGTGTGAGCGGCTGTCTGGGAGA 2640
Db
3279 CCCCCAGCAGGCTGAGAGATCAGAGAGGAGTCAAGAGTGTGAGCGGCTGTCTGGGAGA 3338
QY
2641 GGCTGCGGTGGGACGGAAGCGCAAGTGGTCTCTGCGCCCTCAGCAGGCTCTTCCACTTCCC 2700
Db
3339 GGCTGCGGTGGGACGGAAGCGCAAGTGGTCTCTGCGCCCTCAGCAGGCTCTTCCACTTCCC 3398
QY
2701 TGGGAGGACCTGCGCGCTGGATGAGCTGCGAGCTGGGGCTGCGTGAAGAGTGAAGGC 2760
Db
3399 TGGGAGGACCTGCGCGCTGGATGAGCTGCGAGCTGGGGCTGCGTGAAGAGTGAAGGC 3458
QY
2761 CTCCGTGGAGCACAATCTCCCGGATCTTGAAGGGCAGGCCCGGAGGCTTGGAGAGGAGGG 2820
Db
3459 CTCCGTGGAGCACAATCTCCCGGATCTTGAAGGGCAGGCCCGGAGGCTTGGAGAGGAGGG 3518
QY
2821 GCCCCCAGGAAGAGCAGGCTTGTCTCTTCCGGCTCTCAGGCTCTGAAGAGCTGGGA 2880
Db
3519 GCCCCCAGGAAGAGCAGGCTTGTCTCTTCCGGCTCTCAGGCTCTGAAGAGCTGGGA 3578
QY
2881 CCAGGCGCCACATTTCTTAAGGAGGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 2940
Db
3579 CCAGGCGCCACATTTCTTAAGGAGGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 3638
QY
2941 GACACTGGCTGCCAGGTGTACGCCACGAGCTGCCCCAGGCCACCTTGGGACGAAGAGCG 3000
Db
3639 GACACTGGCTGCCAGGTGTACGCCACGAGCTGCCCCAGGCCACCTTGGGACGAAGAGCG 3698
|||||

Qy 3001 A 3001
Db 3699 A 3699

RESULT 5
US-10-697-263-3
; Sequence 3, Application US/10697263
; Patent No. 6812014
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,263
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-697-263-3

Query Match 99.9%; Score 2997.8; DB 4; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTACGGAGCGAATCTGGGCCCCAGGATACAGGGAGCGAGACATCTCTGGCCGCGCTGAG 60
Db 699 CCTACGGAGCGAATCTGGGCCCCAGGATACAGGGAGCGAGACATCTCTGGCCGCGCTGAG 758

Qy 61 CCACCCGCTGTACAGGGGCTGTGGACCACTTTGAGACCGGCAAGACCTCATCTCTCAT 120
Db 759 CCACCCGCTGTACAGGGGCTGTGGACCACTTTGAGACCGGCAAGACCTCATCTCTCAT 818

Qy 121 CTTGGAGCTGTCTATCCGAGGAGCTGTGGACCGCTGTACAGGAGGCGTGTGTGAC 180
Db 819 CTTGGAGCTGTCTATCCGAGGAGCTGTGGACCGCTGTACAGGAGGCGTGTGTGAC 878

Qy 181 GGAGGCGAGGTGAAGGTCTACATCCAGAGCTGTGGAGGGGCTGCACTACCTGCAAG 240
Db 879 GGAGGCGAGGTGAAGGTCTACATCCAGAGCTGTGGAGGGGCTGCACTACCTGCAAG 938

Qy 241 CCATGGGCTTCTCCACCTGGACATAAAGCCCTCTAAACATCTGATGTGTGATCTGCCCCG 300
Db 939 CCATGGGCTTCTCCACCTGGACATAAAGCCCTCTAAACATCTGATGTGTGATCTGCCCCG 998

Qy 301 GGAAGACATTAATAATCTGCGACTTTGGCTTTGGCCAGAAACATCAACCCAGCAGAGCTGCA 360
Db 999 GGAAGACATTAATAATCTGCGACTTTGGCTTTGGCCAGAAACATCAACCCAGCAGAGCTGCA 1058

Qy 361 GTTCAGGAGTACGGCTCCCTGTAGTTGGCTCTCCCGAGATCATTCAGAGAAACCTGT 420
Db 1059 GTTCAGGAGTACGGCTCCCTGTAGTTGGCTCTCCCGAGATCATTCAGAGAAACCTGT 1118

Qy 421 GAGGAGAGCTCCGACATTTGGGCGATGGGTGTGATCTCTACCTAGCGCTGACCTGCTC 480
Db 1119 GAGGAGAGCTCCGACATTTGGGCGATGGGTGTGATCTCTACCTAGCGCTGACCTGCTC 1178

Qy 481 ATCCCCATTTGCCGCGAGAGTGAACGTTGCCACCTCTCTGAACGTTCTGGAGGGGCGCGT 540
Db 1179 ATCCCCATTTGCCGCGAGAGTGAACGTTGCCACCTCTCTGAACGTTCTGGAGGGGCGCGT 1238

Qy 541 GTCATGAGCAGGCCCCATGGCTGCCCACTCAGCGAAGACGCCAAGACTTCATCAAGGC 600

Db 1239 GTCATGAGCAGGCCCCATGGCTGCCACCTCAGCGAAGACGCCAAGACTTCATCAAGGC 1298
Qy 601 TAGCTGTCAGAGAGCCCTCAGGCCGCGCTAGTGGGCGCCAGTGCCTCTCTCCACCCCTG 660
Db 1299 TAGCTGTCAGAGAGCCCTCAGGCCGCGCTAGTGGGCGCCAGTGCCTCTCTCCACCCCTG 1358
Qy 661 GTTCTGAAATCCATGCTGCGGAGGAGGCCACTTTCATCAACCAAGCAGCTCAAGTT 720
Db 1359 GTTCTGAAATCCATGCTGCGGAGGAGGCCACTTTCATCAACCAAGCAGCTCAAGTT 1418
Qy 721 CTTCTGGGCGGAGTGCCTGGCAGCGTTCCCTGATGAGCTACAGTCCATCTCTGTTGAT 780
Db 1419 CTTCTGGGCGGAGTGCCTGGCAGCGTTCCCTGATGAGCTACAGTCCATCTCTGTTGAT 1478
Qy 781 GCGCTCCATCCCTGAGCTGCTGGGGGCCACCCGACAGCCCTCCCTCGGCGTAGCCCCG 840
Db 1479 GCGCTCCATCCCTGAGCTGCTGGGGGCCACCCGACAGCCCTCCCTCGGCGTAGCCCCG 1538
Qy 841 GCACCTTCGAGGAGACCTGGTGGCTCTCAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db 1539 GCACCTTCGAGGAGACCTGGTGGCTCTCAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1598
Qy 901 CGCCCATTTGCCGGGCTAAAGTCACTGCCACCTCTCCCGGTGACACACTCACCACCTGCT 960
Db 1599 CGCCCATTTGCCGGGCTAAAGTCACTGCCACCTCTCCCGGTGACACACTCACCACCTGCT 1658
Qy 961 GCACCCCGGGGCTTCTCTGGGCGCTCGGCGAGCTGAGGAGCCGAGGCGAGTGA 1020
Db 1659 GCACCCCGGGGCTTCTCTGGGCGCTCGGCGAGCTGAGGAGCCGAGGCGAGTGA 1718
Qy 1021 GCCTCTCACCGAGGCCCTAGCTCCGCTGCTCTCCGAGGGTGCAGGCGCCGCGCGC 1080
Db 1719 GCCTCTCACCGAGGCCCTAGCTCCGCTGCTCTCCGAGGGTGCAGGCGCCGCGCGC 1778
Qy 1081 CAGGGCTGGCTCCCGGCGACAGCTCATCCGAGGCTGTTCTACCAACAGGGGGTGA 1140
Db 1779 CAGGGCTGGCTCCCGGCGACAGCTCATCCGAGGCTGTTCTACCAACAGGGGGTGA 1838
Qy 1141 GAGCCCTGAGCACGGGCGCTGGCCCGGGGAGCAGCGCGCACCCGCGCGGGCGGCGCA 1200
Db 1839 GAGCCCTGAGCACGGGCGCTGGCCCGGGGAGCAGCGCGCACCCGCGCGGGCGGCGCA 1898
Qy 1201 CTTGCTGAAGGGCGGGTACATTCGGGGGGCTGCGAGGCTTGGCGAGCACTGATGGA 1260
Db 1899 CTTGCTGAAGGGCGGGTACATTCGGGGGGCTGCGAGGCTTGGCGAGCACTGATGGA 1958
Qy 1261 GCACCGGCTGCTGGAGGAGGAGCGCCAGGGAGGAGCAGGCCACCTCTCTGCGCAAGC 1320
Db 1959 GCACCGGCTGCTGGAGGAGGAGCGCCAGGGAGGAGCAGGCCACCTCTCTGCGCAAGC 2018
Qy 1321 CCCTCATTTGAGACTGCCCTCCGGTGTCTGCTCTGCGCACCCACTTTGGCCCCCTGGGCCA 1380
Db 2019 CCCTCATTTGAGACTGCCCTCCGGTGTCTGCTCTGCGCACCCACTTTGGCCCCCTGGGCCA 2078
Qy 1381 CAGCCACTCCCTGGAAACATGACTCTCGAGCAACCCCGCCCCCTCTCTCGAGGCGCTCGCG 1440
Db 2079 CAGCCACTCCCTGGAAACATGACTCTCGAGCAACCCCGCCCCCTCTCTCGAGGCGCTCGCG 2138
Qy 1441 TGAGGACACAGCACTGCTTACGCCCTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1500
Db 2139 TGAGGACACAGCACTGCTTACGCCCTCCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2198
Qy 1501 CCCTCAGGGCTCCCAAGCAGCTTCCATCCACTGCTGGCGCACCCAGGCACTCTCAGCCAGA 1560
Db 2199 CCCTCAGGGCTCCCAAGCAGCTTCCATCCACTGCTGGCGCACCCAGGCACTCTCAGCCAGA 2258
Qy 1561 GAGGCCATCCCGGAGCAGCCCTTTGGGGGAGAGCAGCCCTTTCTGCCACCCCAAGCAGGG 1620
Db 2259 GAGGCCATCCCGGAGCAGCCCTTTGGGGGAGAGCAGCCCTTTCTGCCACCCCAAGCAGGG 2318
Qy 1621 TTCTGCCCCCAGGAGGGCTGAGCCGCCACCCAGCAGGTTGCCCGCCATGCGCTCTCTGCTC 1680

Db 2319 TTCTGCCCCCAGGAGGCTGACGCCCCACCCAGCAGTTGTCCTCCATGCTCTGGCTC 2378
Qy 1681 CTTCCCTCCAGGATCTTGAAGAGGCCCCCTTAGTACCCTCAAGCCCTCTTGGGACA 1740
Db 2379 CTTCCCTCCAGGATCTTGAAGAGGCCCCCTTAGTACCCTCAAGCCCTCTTGGGACA 2438
Qy 1741 GCCCAGGACCCCTGCTCCCTGCAAGAGAGCCGCCCATTTGGACTCTAAGATGGGGCC 1800
Db 2439 GCCCAGGACCCCTGCTCCCTGCAAGAGAGCCGCCCATTTGGACTCTAAGATGGGGCC 2498
Qy 1801 TGGAGACATCTCTTCTTGGAGGCCCCAAACCCGGGCCCTGCAAGTTCCCCAGGGTCAGC 1860
Db 2499 TGGAGACATCTCTTCTTGGAGGCCCCAAACCCGGGCCCTGCAAGTTCCCCAGGGTCAGC 2558
Qy 1861 CTCCAGGCGAGCTCTTCCCAAGTAGCTCCTCAGGGTGGGCTCTCCCAAGTGGGAC 1920
Db 2559 CTCCAGGCGAGCTCTTCCCAAGTAGCTCCTCAGGGTGGGCTCTCCCAAGTGGGAC 2618
Qy 1921 AGAGCCTGGGCCCTCTGATGCGAGGGCTGACCCAGGAGCTGAGGATCTGCCGA 1980
Db 2619 AGAGCCTGGGCCCTCTGATGCGAGGGCTGACCCAGGAGCTGAGGATCTGCCGA 2678
Qy 1981 CTCACACCCACCTTGACGGGCTCTAGGAACAGGCGACCATGCGCAAGTTCTCCCTGGG 2040
Db 2679 CTCACACCCACCTTGACGGGCTCTAGGAACAGGCGACCATGCGCAAGTTCTCCCTGGG 2738
Qy 2041 TGGTGGGGGGCTACGAGGCGTGGCTGGCTATGAGCACTTTGCCCTTTGGTGGAGATGC 2100
Db 2739 TGGTGGGGGGCTACGAGGCGTGGCTGGCTATGAGCACTTTGCCCTTTGGTGGAGATGC 2798
Qy 2101 AGGGGGATGCTGGGGCAGGGGCCCATGTGGGCGAGGATAGCTGGGCTGTGCCAGTC 2160
Db 2799 AGGGGGATGCTGGGGCAGGGGCCCATGTGGGCGAGGATAGCTGGGCTGTGCCAGTC 2858
Qy 2161 GGAGGAGGAGGAGGAGGAGGCGGCGAGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2220
Db 2859 GGAGGAGGAGGAGGAGGAGGCGGCGAGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2918
Qy 2221 CAGGGCTGAGAGCCACTGCCCCAGGTCAGTGCAGAGGCTGTGCTGAGGTCGGCAGGGC 2280
Db 2919 CAGGGCTGAGAGCCACTGCCCCAGGTCAGTGCAGAGGCTGTGCTGAGGTCGGCAGGGC 2978
Qy 2281 TCCACACGAGGAGCTCCAGAGGCCACCCATGCGAGGAGCATCGGGAGGTCCTCCCTGGT 2340
Db 2979 TCCACACGAGGAGCTCCAGAGGCCACCCATGCGAGGAGCATCGGGAGGTCCTCCCTGGT 3038
Qy 2341 GCAGATCGGGACCTGTGAGGTGATGCGAGGCGGCGACACAATATCCCTGGACATTTTC 2400
Db 3039 GCAGATCGGGACCTGTGAGGTGATGCGAGGCGGCGGCGACACAATATCCCTGGACATTTTC 3098
Qy 2401 CGAGGTGGACCCCGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
Db 3099 CGAGGTGGACCCCGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 3158
Qy 2461 CGAGTTTATGATCTTCAGGAAGTCCCAAGTCCGCTCAGGCGAGCGGCTCCCGCAT 2520
Db 3159 CGAGTTTATGATCTTCAGGAAGTCCCAAGTCCGCTCAGGCGAGCGGCTCCCGCAT 3218
Qy 2521 GGCTGAGGAGGAGCTGGCCGAGTTCCCGAGGCCACAGTGGCCCTGGCCAGGTGAACCTGGG 2580
Db 3219 GGCTGAGGAGGAGCTGGCCGAGTTCCCGAGGCCACAGTGGCCCTGGCCAGGTGAACCTGGG 3278
Qy 2581 CCCCCAGCGAGGCTGGAGATCACAGAGGAGTCAGAGGATGTGGACGCGCTGTCTGGCAGA 2640
Db 3279 CCCCCAGCGAGGCTGGAGATCACAGAGGAGTCAGAGGATGTGGACGCGCTGTCTGGCAGA 3338
Qy 2641 GGCTGCCGTGGGAGGAGCGCAAGTGGTCTCGCCGTTCAGCGAGCTCTTCCACTTCCC 2700
Db 3339 GGCTGCCGTGGGAGGAGCGCAAGTGGTCTCGCCGTTCAGCGAGCTCTTCCACTTCCC 3398
Qy 2701 TGGGAGGACCTCGCGCTGGATAGCTTGCAGAGCTGGGGCTGGTGGAGAGTGAAGGC 2760
Db 3399 TGGGAGGACCTCGCGCTGGATAGCTTGCAGAGCTGGGGCTGGTGGAGAGTGAAGGC 3458

Qy 2761 CTCGTCGAGCAGCATCTCCGGATCTCTGAAGGCGAGCCCGAGGTCTGGAGAGGAGGG 2820
Db 3459 CTCGTCGAGCAGCATCTCCGGATCTCTGAAGGCGAGCCCGAGGTCTGGAGAGGAGGG 3518
Qy 2821 GCCCCCCAGGAAGAAGCAGGCGCTTGTCTCTCCGGCTCTCAGGTCTGAAGAGCTGGGA 2880
Db 3519 GCCCCCCAGGAAGAAGCAGGCGCTTGTCTCTCCGGCTCTCAGGTCTGAAGAGCTGGGA 3578
Qy 2881 CCGAGCCCCGACATCTCTAAGGAGGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 2940
Db 3579 CCGAGCCCCGACATCTCTAAGGAGGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCAGT 3638
Qy 2941 GACACTGGCCTGCGAGGTGTACGCCAGCAGCTGCCAGGCGACCTGGAGCAAGAGCG 3000
Db 3639 GACACTGGCCTGCGAGGTGTACGCCAGCAGCTGCCAGGCGACCTGGAGCAAGAGCG 3698
Qy 3001 A 3001
Db 3699 A 3699

RESULT 6

US-09-159-385-4
; Sequence 4, Application US/09159385
; Patent No. 5958748
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: PH-569
; CURRENT APPLICATION NUMBER: US/09/159,385
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10)..(1353)
US-09-159-385-4

Query Match 5.7%; Score 172; DB 2; Length 1429;

Best Local Similarity 54.4%; Pred. No. 5.1e-27;

Matches 371; Conservative 0; Mismatches 305; Indels 6; Gaps 1;

Qy 17 CGGCGCCAGGCATACAGGAGCGAGACATCTGTGCGCGCTGAGCCACCGCTGGTCAG 76
Db 181 CGGAGGAGATCGAAGCGAGGTGAGCATCTGCGCGAGATCCGCCACCCCAACATCATA 240
Qy 77 GGGCTCTGAGCAGTTTGAGACCCGCAAGACCTCATCTCATCTCTGGAGCTGTGTCA 136
Db 241 ACATGTCATGACGTGTTTCGAGAAACAGACAGATGTGCTGTGATCTCTGGAGCTGGTGTCC 300
Qy 137 TCCGAGGAGCTGTGGACCCGCTGTACAGGAAGGCGTGTGACGAGGCGGAGCTCAAG 196
Db 301 GGTGGCGAGCTTTCGACTTCTTGGCCGAGAGAGGTCATTGACGGAGGATGAGGCCACG 360
Qy 197 GTCTACATCCAGCAGCTGTGGAGGGCTGCATCTGCTGACACAGCCATGGCGTTCTCCAC 256
Db 361 CAGTTCCTCAAAACAAATCTTAGACGGTGTCCACTACCTGCACTCAAGCGCATCGCACAC 420
Qy 257 CTGGACATAAAGCCCTTAACATCTCTGATGTT-----GCATCTGTGCGCGGAAGACATT 310
Db 421 TTTGACCTGAAGCCCGAGAACATCATGTTGTCGACAAGCAGCAGCGCCAGCCCGGCATT 480
Qy 311 AATATCTGCACCTTTGGCTTTGCCAGAACATCACCCAGCAGAGCTGCAGTTTCAGCCAG 370
Db 481 AAGCTCATGACTTTTGCAATCGCGCAAGGATCGAGGCTGGCAGGAGTTCAGAACATC 540

Qy	371	TACGGCTCCCTGAGTTTCGTCTCCCCGAGATCATCCAGCAGAAACCTGTGAGCGAAGCC	430
Db	541	TTTGGCACACCCGAGTTTGTTCGCCCCGAGATCGTGAATATGAGCACTTGGCTTGGAG	600
Qy	431	TCCGACATTTGGGGCATGGGTGTATCTCTTACCTCAGCTGACCTGCTCATCCCATTT	490
Db	601	GCTGACATGTGGAGCATTTGGGGTTCATCACCTACATCTCTGAGGGGAGCGTCCCATTC	660
Qy	491	GCCGGCGAGAGTGCAGGTGCCACCTCTCTGAACGTCTGGAGGGGGCGGTGATGGAGC	550
Db	661	CTGGCGGAGACCAAGCAGGAGACGCTGACAAATCTCAGCAGTGAATATGACTTTGAT	720
Qy	551	AGCCCCATGGCTGCCCTCAGCGAAGACGCCAAAGACTTCATCAAGCTACGCTGCAG	610
Db	721	GAGGAATACTTCAGGAGACACAGCGAGCTGGCCAAAGGACTTCATCGCAGGCTGCTGGTC	780
Qy	611	AGAGCCCTTCAGGCCCGGCTTAGTCGGGCCCAGTGCCTCTCCACCCCTGGTTCCTGAAA	670
Db	781	AAAGACCCCAAGAGGAGATGACCATCGCACAGAGCCTGGAGCATTCCTGGATCAAGGTG	840
Qy	671	TCCATGCTGCGGAGGAGGCC	692
Db	841	CGCAGGGCGGAGGACCGCGCCC	862

RESULTS

```

US-09-186-277-4
; RESOLUTION:
; Patent No. 6171841
; Application US/09186277
;
; GENERAL INFORMATION:
;
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
;
; TITLE OF INVENTION: DNA CODING FOR SE
;
; FILE REFERENCE: 081356/0128
; CURRENT APPLICATION NUMBER: US/09/186
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: JP97/2615
; EARLIER FILING DATE: 1997-09-26
;
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 4
;
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: Mus musculus
;
; FEATURE:
;
; NAME/KEY: CDS
; LOCATION: (10)..(1353)
;
; US-09-186-277-4

```

QY	311	AAATCTCGCATTTGGCTTTGGCCGAGAACATCATCCCGAGCAGAGCTGCAGTTTCAGCCAG	370
DB	481	AAGCTCATCGACTTTGGCATCGCGACAGGATCGAGGCTGGCAGCGAGTTCAAGAAATCATC	540
QY	371	TACGGCTCCCTGAGTTTCGTCTCCCCCGAGATCATCCAGCAGAAACCTGTGAGCGAAGCC	430
DB	541	TTTGCGACACCCGAGTTTGTGCGCCCCGAGATCGTGAATATGAGCCACTTGGCTTGGAG	600
QY	431	TCCGACATTTGGGCGCATGGGTGTTCATCTCTTACCTCAGCCTCAGCTGCTCATCCCCCATTT	490
DB	601	GCTGNCATGTGGAGCATTTGGCGTCAATACCTATACCTCTCTGAGCGGAGCGTCCCCATTC	660
QY	491	GCCGCGAGATGACCGGTGCACCCCTCTGTAAACGTCTGGAGGGCGGTGTCATGGAGC	550
DB	661	CTGGGCGGAGACCAAGCAGGAGACGCTGACGAACATCTCAGCAGTGAATATGACTTTGAT	720
QY	551	AGCCCCATGGTGCCTCCACTCAGCGAAGACGCCAAGACTTCATCAAGGCTACGCTGCAG	610
DB	721	GAGGAATACTTTCAGCAGACCAAGCGAGCTGGCGCAAGGACTTCATCCGAGGCTGCTGTGTC	780
QY	611	AGAGCCCCCTCAGGCGCGGCGCTAGTGCGGCGCCAGTGCCTCTCCCAACCCCTGGTTCTCTGAAA	670
DB	781	AAAGACCCCAAGAGGAGGATGACCATCGCAGAGCCTGGAGCATCTCTGGATCAAGGTG	840
QY	671	TCCATGCTCGGAGGAGGCGCC	692
DB	841	CGCAGGCGCGAGGACCGGCGCCC	862

RESULT 8

```

US-09-949-016-1966
; Sequence 1966, Application US/09949A016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN
; FILE REFERENCE: WITH HUMAN DISEASE,
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1966
; LENGTH: 2157
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1966

```

197 GTCTACATCCAGAGCTGGTGGAGGGGCTGCACCTACCTGCACAGGCATGGCGTTCTCCAC 256

Db 497 CAGTTCTCAAGCAGATCCTCGACGGGTTTCACTACCTGCACTCTAAGCGCATCGCAC 556
Qy 257 CTGACATATAAGCCCTCTAAATCCTGATGCTGATGCTGCTGCGGGAAGAC-----ATT 310
Db 557 TTTGACCTGAAGCCGGAATAATCATGCTGCTGGAAGAAGACGTCGCCAACCCACGAATC 616
Qy 311 AAAATCTGGCACTTTGGCTTTGGCCAGAAATCAACCCAGCAGAGCTGCAGTTCAAGCCAG 370
Db 617 AAGCTCATCGACTTCGGCATCGCGCAAGAATCAGGGGGGAACGAGTTCAAGAACATC 676
Qy 371 TACGGCTCCCTGAGTTTGGTCTCCCGGAGATCATCCAGCAAGAACCCCTGTGAGGGAAGCC 430
Db 677 TTCGGCACCCCGGAGTTTGGCCCGCAGAGATTGTGAATATGAGCGCTGGGCTGGAG 736
Qy 431 TCCGACATTTGGGCGCATGGGTGTCATCTCCCTACCTCAGCCTGACCTGCTCATCCCATTT 490
Db 737 GCGGACATGTGGAGCATCGGTGTCTATCACCTATATCTCTCTGAGCGGTGCATCCCGTTC 796
Qy 491 GCCGGCAGAGTACCGTGCACACCTCTCTGAACGTCCTGGAGGGGCGGTGTCTATGGAGC 550
Db 797 CTGGCGGAGCAAGCAGGAGACGCTTACCAACATCTCAGCGGTGAATCAGACTTCGAC 856
Qy 551 AGCCCCCTGCTGCCACCTCAGGGAAGACGCCAAAGATTTTCATCAAGGCTTACGCTGCAG 610
Db 857 GAGGAGTACTTCAGCAACACCAGCGAGCTGSCCAAGGACTTTCATTTCGCGGCTGCTGTC 916
Qy 611 AGAGCCCTCAGGCGCGGCTAGTGGCGCCAGTGCCTCTCCACCCCTGTTT 663
Db 917 AAAGATCCCAAGCGGAGAAATGACCATTGCCCGAGAGCCTGGAACATTCCTGGAT 969

RESULT 9

US-09-949-016-195
; Sequence 195, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 195
; LENGTH: 2105
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-195

Query Match 5.6%; Score 168.6; DB 4; Length 2105;
Best Local Similarity 54.8%; Pred. No. 3e-26;
Matches 358; Conservative 0; Mismatches 289; Indels 6; Gaps 1;

Qy 17 CGGGCCAGGATACAGGAGCGGAGATCCTGGCCGCGCTGAGCCACCCGCTGGTCAAG 76
Db 265 CGGGAGGAGATCGAGCGGAGGTGAACATCTCTGGGAGATCCGGCACCCCAACATCATC 324
Qy 77 GGGCTGTGGACCACTTTGAGACCGCAAGACCTCATCTCTCTGAGCTGTGCTCA 136
Db 325 ACCCTGACGACATCTTCGAGAACAGACGAGCTGTCTCTCATCTCTGAGCTGTCTCT 384
Qy 137 TCCGAGGAGCTGTGGACCGCTGTACAGGAAGGGCGTGTGACGAGGCGGAGGTCAAG 196
Db 385 GCGGGGAGCTCTTTGACTTCTCTGGCGGAGAAAGAGTCGCTGACGGAGGACGAGGCCACC 444
Qy 197 GTCTACATCCAGCAGCTGGTGGAGGGGCTGCATCTACCTGCAAGCATGGCGGTCTCCAC 256

Db 445 CAGTTCTCAAGCAGATCCTGGACGGGTTTCACTACCTGCACTCTAAGCGCATCGCAC 504
Qy 257 CTGACATATAAGCCCTCTAAATCCTGATGCTGATGCTGCTGCGGGAAGAC-----ATT 310
Db 505 TTTGACCTGAAGCCGGAATAATCATGCTGCTGGAAGAAGACGTCGCCAACCCACGAATC 564
Qy 311 AAAATCTGGCACTTTGGCTTTGGCCAGAAATCAACCCAGCAGAGCTGCAGTTCAAGCCAG 370
Db 565 AAGCTCATCGACTTCGGCATCGCGCAAGAATCAGGGGGGAACGAGTTCAAGAACATC 624
Qy 371 TACGGCTCCCTGAGTTTGGTCTCCCGGAGATCATCCAGCAAGAACCCCTGTGAGGGAAGCC 430
Db 625 TTCGGCACCCCGGAGTTTGGCCCGCAGAGATTGTGAATATGAGCGCTGGGCTGGAG 684
Qy 431 TCCGACATTTGGGCGCATGGGTGTCATCTCCCTACCTCAGCCTGACCTGCTCATCCCATTT 490
Db 685 GCGGACATGTGGAGCATCGGTGTCTATCACCTATATCTCTCTGAGCGGTGCATCCCGTTC 744
Qy 491 GCCGGCAGAGTACCGTGCACACCTCTCTGAACGTCCTGGAGGGGCGGTGTCTATGGAGC 550
Db 745 CTGGCGGAGCAAGCAGGAGACGCTTACCAACATCTCAGCGGTGAATCAGACTTCGAC 804
Qy 551 AGCCCCCTGCTGCCACCTCAGGGAAGACGCCAAAGATTTTCATCAAGGCTTACGCTGCAG 610
Db 805 GAGGAGTACTTCAGCAACACCAGCGAGCTGSCCAAGGACTTTCATTTCGCGGCTGCTGTC 864
Qy 611 AGAGCCCTCAGGCGCGGCTAGTGGCGCCAGTGCCTCTCCACCCCTGTTT 663
Db 865 AAAGATCCCAAGCGGAGAAATGACCATTGCCCGAGAGCCTGGAACATTCCTGGAT 917

RESULT 10

US-09-159-385-3
; Sequence 3, Application US/09159385
; Patent No. 5958748
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: PH-569
; CURRENT APPLICATION NUMBER: US/09/159,385
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)..(1455)
US-09-159-385-3

Query Match 5.6%; Score 168.6; DB 2; Length 2132;
Best Local Similarity 54.8%; Pred. No. 3e-26;
Matches 358; Conservative 0; Mismatches 289; Indels 6; Gaps 1;

Qy 17 CGGGCCAGGATACAGGAGCGGAGATCCTGGCCGCGCTGAGCCACCCGCTGGTCAAG 76
Db 265 CGGGAGGAGATCGAGCGGAGGTGAACATCTCTGGGAGATCCGGCACCCCAACATCATC 324
Qy 77 GGGCTGTGGACCACTTTGAGACCGCAAGACCTCATCTCTCTGAGCTGTGCTCA 136
Db 325 ACCCTGACGACATCTTCGAGAACAGACGAGCTGTCTCTCATCTCTGAGCTGTCTCT 384
Qy 137 TCCGAGGAGCTGTGGACCGCTGTACAGGAAGGGCGTGTGACGAGGCGGAGGTCAAG 196
Db 385 GCGGGGAGCTCTTTGACTTCTCTGGCGGAGAAAGAGTCGCTGACGGAGGACGAGGCCACC 444
Qy 197 GTCTACATCCAGCAGCTGGTGGAGGGGCTGCATCTACCTGCAAGCATGGCGGTCTCCAC 256

Db 445 CAGTTCCTCAAGCAGATCCCTGGACGGGCTTCACTACCTGACACTCTAAGCGCATGCGACAC 504
Qy 257 CTGGACATAAAGCCCTCTAAACATCCTGATGGTGATCCTGCGCCGGGAAGAC-----ATT 310
Db 505 TTTGACCTGAAGCCGGAANAACATCATCTGCTGCAAGAAGCTGCCAACCCACGAATC 564
Qy 311 AAAATCTGCGACTTTGGCTTTGCCCCAGAAATCATCAACCCAGAGAGCTGCAATTCAGCCAG 370
Db 565 AAGCTCATCGACTTCGGCATCGCGCACAAAGATCGAGGCGGGAAACGAGTTTCAAGAAATC 624
Qy 371 TAGCGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGAGACCTGTGAGGAGGCC 430
Db 625 TTCGGCACCCCGGAGTTTGGCCCCAGAGATTGTGAATATGAGCCGCTGGGCTGGAG 684
Qy 431 TCCGACATTTGGGCCATAGGGTGTCTCTACCTCAGCCTGACCTGCTCATCCCCATT 490
Db 685 GCGGACATGTGGACATCGGTGTATCACTATATCTCTGAGCGGTGATCCCCGTTT 744
Qy 491 GCCGGCAGAGTGACCGTGCACCCCTCTGAACGTCTGGAGGGGCGCGTGTATGAGC 550
Db 745 CTGGCGAGACCAAGCAGGAGACGCTCACCAATCTCAGCCGTGAATACGACTTCGAC 804
Qy 551 AGCCCCATGGTGCCTCAGCAAGACGCGCAAGAGCTTCATCAAGGCTACGCTGAC 610
Db 805 GAGGAGTACTTCAGCAACACCGAGCGAGCTGGCCAAAGGACTTCAATTCGCGGCTCTCGTC 864
Qy 611 AGAGCCCTCAGGCCCGGCTAGTGGCGCCAGTGCCTCTCCACCCCTGGTT 663
Db 865 AAAGATCCCAAGCGGAGATGACCATTTGCCAGAGCCTGGAACTTCTTGAT 917

RESULT 11

US-09-186-277-3
; Sequence 3, Application US/09186277
; Patent No. 6171841
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: 081356/0128
; CURRENT APPLICATION NUMBER: US/09/186,277
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)..(1455)
US-09-186-277-3

Query Match 5.6%; Score 168.6; DB 3; Length 2132;
Best Local Similarity 54.8%; Pred. No. 3e-26;
Matches 358; Conservative 0; Mismatches 289; Indels 6; Gaps 1;
Qy 17 CGGCCCGAGGATATACAGGAGCGAGACATCTGCGCGCGGTGAGCCACCCGCTGGTCAAG 76
Db 265 CGGAGGAGATCGAGCGGAGGTGAACATCTCGGGGAGATCCGGCACCCCAACATATC 324
Qy 77 GGGCTGTGACACAGTTTGAGACCGGACCGACCTCATCTCTGAGCTGTGCTCA 136
Db 325 ACCGTGACGACATCTTCGAGAAACAGACGACGTGGTCTCTATCTGAGAGCTGCTCT 384
Qy 137 TCCGAGAGGTGCTGGACCGCTGTACAGGAAGGGCGTGGTGAACGAGGCGGAGGTCAAG 196
Db 385 GCGGGGAGCTCTTTGACTTCTTGGCGGAGAAAGAGTGCCTGACGGAGGACGAGGCCACC 444
Qy 197 GTCTACATCCAGCAGCTGGTGGAGGGGCTGCACTACCTGCACAGCCATGGCGTTCTCCAC 256

Db 445 CAGTTCCTCAAGCAGATCCCTGGACGGGTTCACTACCTGACACTCTAAGCGCATGCGACAC 504
Qy 257 CTGGACATAAAGCCCTCTAAACATCCTGATGGTGATCCTGCGCCGGGAAGAC-----ATT 310
Db 505 TTTGACCTGAAGCCGGAANAACATCATCTGCTGCAAGAAGCTGCCAACCCACGAATC 564
Qy 311 AAAATCTGCGACTTTGGCTTTGCCCCAGAAATCATCAACCCAGAGAGCTGCAATTCAGCCAG 370
Db 565 AAGCTCATCGACTTCGGCATCGCGCACAAAGATCGAGGCGGGAAACGAGTTTCAAGAAATC 624
Qy 371 TAGCGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGAGACCTGTGAGGAGGCC 430
Db 625 TTCGGCACCCCGGAGTTTGGCCCCAGAGATTGTGAATATGAGCCGCTGGGCTGGAG 684
Qy 431 TCCGACATTTGGGCCATAGGGTGTCTCTACCTCAGCCTGACCTGCTCATCCCCATT 490
Db 685 GCGGACATGTGGACATCGGTGTATCACTATATCTCTGAGCGGTGATCCCCGTTT 744
Qy 491 GCCGGCAGAGTGACCGTGCACCCCTCTGAACGTCTGGAGGGGCGCGTGTATGAGC 550
Db 745 CTGGCGAGACCAAGCAGGAGACGCTCACCAATCTCAGCCGTGAATACGACTTCGAC 804
Qy 551 AGCCCCATGGTGCCTCAGCAAGACGCGCAAGAGCTTCATCAAGGCTACGCTGAC 610
Db 805 GAGGAGTACTTCAGCAACACCGAGCGAGCTGGCCAAAGGACTTCAATTCGCGGCTCTCGTC 864
Qy 611 AGAGCCCTCAGGCCCGGCTAGTGGCGCCAGTGCCTCTCCACCCCTGGTT 663
Db 865 AAAGATCCCAAGCGGAGATGACCATTTGCCAGAGCCTGGAACTTCTTGAT 917

RESULT 12

US-09-949-016-5449
; Sequence 5449, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PastSeq for Windows Version 4.0
; SEQ ID NO 5449
; LENGTH: 1847
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-5449

Query Match 5.0%; Score 148.6; DB 4; Length 1847;
Best Local Similarity 51.5%; Pred. No. 5e-22;
Matches 368; Conservative 0; Mismatches 144; Indels 3; Gaps 1;
Qy 34 GGAGCGAGACATCCTGGCGCGCTGAGCCACCCGCTGGTCAACGGGGTGTGACCAAGTT 93
Db 874 GGAGATTAGCATCATGAATGCTCCTCCACCCACCTAAAGCTGGTCCAGTGTGTGATGCTT 933
Qy 94 TGAGACCCCGCAAGACCTCATCTCTGAGCTGTGCTCATCCGAGGAGCTGCTGGA 153
Db 934 TGAAGAAAAGGCCAAACATGTCATGCTTGGTGTGAGATCGTGTGAGGGGAGCTGTTGA 993
Qy 154 CCG-----CCTGTACAGGAAGGGCGTGTGACGAGGCGGAGGTCAAGGTCTACATCCAGCA 210
Db 994 GCGCATCATTGACAGGAGACTTTTGAGCTGACGGAGCGGTGATGATCAAGTACATGCGGCA 1053


```
Db 1107 GAGGATTGTGGATGAGGACTACCATCTGACCGAGGTGGACACCATGTGTTGTGTCAGGCA 1166
Qy 211 GCTGGTGGAGGGGTGACACTACCTGCGACAGCCATGGGGTCTTCCACCTGGACATAAAGCC 270
Db 1167 GATCTGTGACGGGATCCTCTTTCATGSCAACAGATGAGGGTTTTGACACCTGGACCTCAAGCC 1226
Qy 271 CTCTAAACATCTGATGTGTCATCTGCGCGGAAGACATTAAATCTGCGACTTTGGCTT 330
Db 1227 AGAGAACATCTGTGTGTCAACACACCGGSCATTTGGTGAAGATCAATTGACTTTGGCCT 1286
Qy 331 TGCCGAGAACATACCCGACGAGCTGCAAGTTGAGCCAGTACGGCTCCCTCGAGTTTCGT 390
Db 1287 GGCACGGAGGTATAACCCCAACGAGAGAGCTGAAGGTGAACCTTTGGGACCCCGAGAGTTCT 1346
Qy 391 CTCCTCCGAGATCATCAGCAGAACCTGTGAGGGAAGCTCCGACATTTGGGCCATGGG 450
Db 1347 GTCACTGAGGTGGTGAATTATGACCAATCTCCGATAAGACAGACATGTGGAGTATGGG 1406
Qy 451 TGTCACTCTCTACCTGAGCCTGACCTGCTCATCTCCCATTTTGGCGGAGAGTGACCGTGC 510
Db 1407 GGTGATACCTACATGTGCTGAGGGGCTCTCCCTTCTCTGGAGATGATGACAGAGA 1466
Qy 511 CACCTCTCTGAACGTCTTGGAGGGCGGTGTCTATGAGCAGAGCCCATGGCTGCCCACT 570
Db 1467 GACCTAAACAACTTCTATCTGGCAACTGGTACTTTGATGAAGAGACCTTTGAGGGCGT 1526
Qy 571 CAGCGAAGACCCAAAGACTTCATCAAGGCTACGCTGAGAGAGCCCTCAGGCGCCGCC 630
Db 1527 ATCAGAGAGGCCAAAGACTTTGTCTCAACCTCATCGTCAAGGACCGAGGGCCCGGAT 1586
Qy 631 TAGTGGCGCCAGTGCCTCTCCACCCCTGGTTCTGTAA 669
Db 1587 GAACGCTGCCAGTGTCTGCGCCATCCCTGCTCAACAA 1625
```

RESULT 15

```
US-09-797-039-7
; Sequence 7, Application US/09797039
; Patent No. 6730491
; GENERAL INFORMATION:
; APPLICANT: Olandt, Peter J.
; TITLE OF INVENTION: 2504, 15977, AND 14760, NOVEL PROTEIN
; TITLE OF INVENTION: KINASE FAMILY MEMBERS AND USES THEREFOR
; FILE REFERENCE: 10448-017001
; CURRENT APPLICATION NUMBER: US/09797,039
; CURRENT FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 2046
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)...(1906)
; NAME/KEY: misc_feature
; LOCATION: (1)...(2046)
; OTHER INFORMATION: n = A,T,C or G
US-09-797-039-7
```

```
Query Match 4.8%; Score 143; DB 4; Length 2046;
Best Local Similarity 52.6%; Pred. No. 7.9e-21;
Matches 336; Conservative 0; Mismatches 300; Indels 3; Gaps 1;

Qy 34 GGAGCGAGACATCTTGGCGGGGTGAGCCACCGCTGGTCAAGGGGTGCTGGACCAAGTT 93
Db 1105 GGAGATTGAGTGTATGAACCAAGCTGAACCAACCGCAATCTGATCCAGCTGATGAGCCAT 1164
Qy 94 TGAGACCGCGCAAGACCCCTCATCTCTCTGAGAGCTGTGCTCATCCGAGGAGCTGCTGGA 153
```

```
Db 1165 CGAGACTCCGATGAGATCGTCTGTGTTATGAGTACATCGAGGGCGGAGAGCTCTTCGA 1224
Qy 154 CCG---CCTGTACAGGAAGGGCGTGTGACGGAGGGCGAGGTCAAGGTCTTACATCCAGCA 210
Db 1225 GAGGATTGTGGATGAGGACTTACCATCTGACCGAGGTGGAACCATGTGTGTTGTCAAGCA 1284
Qy 211 GCTGGTGGAGGGGTGTCACCTACCTGACAGCCATGGCGTTCTCCACCTGGACATAAAGCC 270
Db 1285 GATCTGTGAGGGATCCTCTTTCATGACACAGATGAGGGTTTTGCACCTGGACCTCAAGCC 1344
Qy 271 CTCTAACATCTGATGTGTGATCCTGCTCCCGGGAAGACATTAAATCTGCGACTTTGGCTT 330
Db 1345 AGAGAACATCTGTGTGTCAACACCCACCGGCGCATTTGGTGAAGATCAATTGACTTTGGCCT 1404
Qy 331 TGCCGAGAACATCACCCGACGAGCTGCGAGTTTCAGCCAGTACGGCTCCCTCGAGTTTCGT 390
Db 1405 GGCACGGAGGTATAACCCCAACGAGAGCTGAAGGTGAACCTTTGGGACCCCGAGAGTTCT 1464
Qy 391 CTCCTCCGAGATCATCCAGCAGAACCTGTGAGCGAAGCCTCCGACATTTGGGCCATGGG 450
Db 1465 GTCACTGAGGTGGTGAATTTATGACCAATCTCCGATAAGACAGACATGTGGAGTATGG 1524
Qy 451 TGTCACTCTCTACCTGAGCCTGACCTGCTCATCTCCCATTTTGGCGGAGAGTGACCGTGC 510
Db 1525 GGTGATCACCTACATGCTGTGAGCGGCTCTCCCTTCTCTGGGAGATGATGACACAGA 1584
Qy 511 CACCTCTCTGAACGTCTTGGAGGGCGGTGTCTATGAGCAGAGCCCATGGCTGCCCACT 570
Db 1585 GACCTTAAACAACGTTCTATCTGGCAACTGGTACTTTGTGTAAGAGACCTTTGAGGCCGT 1644
Qy 571 CAGCGAAGAGCGCCAAAGACTTCATCAAGGCTACGCTGACAGAGAGCCCTCAGGCGCCGCC 630
Db 1645 ATCAGAGAGGCCCAAGACTTTGTCTCAACCTCATGTCAGAGACCGAGGGCCCGGAT 1704
Qy 631 TAGTGGCGCCAGTGCCTCTCCACCCCTGGTTCTGTAA 669
Db 1705 GAACGCTGCCAGTGTCTGCGCCATCCCTGGCTCAACAA 1743
```

Search completed: March 21, 2005, 14:39:39

Job time : 533.683 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:02:07 ; Search time 193.228 Seconds
(without alignments)
9323.380 Million cell updates/sec

Title: US-10-077-130-6_COPY_13600_14700
Perfect score: 1101
Sequence: 1 cacagcagcacactgtgac.....agcaggaggcccatgttc 1101

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	45.8	4.2	1350	4	US-09-902-540-5939
C 2	45.8	4.2	1351	4	US-09-902-540-192
C 3	45	4.1	2112	4	US-09-902-540-3277
4	45	4.1	17480	4	US-09-902-540-1151
5	45	4.1	33529	3	US-09-144-085-3
C 6	44.8	4.1	765	4	US-09-252-991A-15050
C 7	44.8	4.1	1329	4	US-09-252-991A-14932
C 8	44.8	4.1	5121	4	US-09-252-991A-15189
9	44	4.0	735	4	US-09-252-991A-14239
C 10	44	4.0	1410	4	US-09-252-991A-14025
C 11	43.8	4.0	1811	4	US-09-902-540-375
C 12	42.6	3.9	1017	4	US-09-252-991A-14630
C 13	42.6	3.9	2502	4	US-09-252-991A-14994
C 14	42.6	3.9	2505	4	US-09-252-991A-14750
C 15	42	3.8	3397	4	US-09-949-016-2188
C 16	42	3.8	119801	4	US-09-949-016-13453
C 17	41.8	3.8	4403765	3	US-09-103-840A-2
C 18	41.8	3.8	4411529	3	US-09-103-840A-1
C 19	41.4	3.8	575	4	US-09-270-767-3998
C 20	41.4	3.8	575	4	US-09-270-767-19280
C 21	41.4	3.8	2010	4	US-09-902-540-9547
C 22	41.4	3.8	14807	4	US-09-902-540-1085
C 23	41.2	3.7	1926	3	US-09-249-585A-2
24	41.2	3.7	1926	3	US-09-410-399-3
25	41.2	3.7	2580	3	US-09-050-863-2
26	41.2	3.7	2580	3	US-09-359-081-2
C 27	41.2	3.7	5452	2	US-09-130-114-1

C 28	41.2	3.7	8705	4	US-09-647-344A-14	Sequence 14, Appl
29	41.2	3.7	9600	3	US-08-910-647-1	Sequence 1, Appl
30	41.2	3.7	9600	3	US-09-620-925-1	Sequence 1, Appl
31	41.2	3.7	10596	1	US-07-884-811-15	Sequence 15, Appl
32	41.2	3.7	10596	1	US-07-885-971-15	Sequence 15, Appl
33	41.2	3.7	10596	1	US-08-087-783A-15	Sequence 15, Appl
34	41.2	3.7	10596	1	US-08-194-088B-15	Sequence 15, Appl
35	41.2	3.7	10596	2	US-08-194-087-15	Sequence 15, Appl
36	41.2	3.7	10596	5	PCT-US93-04648-15	Sequence 15, Appl
C 37	41.2	3.7	16080	4	US-09-724-566A-48	Sequence 48, Appl
C 38	41.2	3.7	16080	4	US-09-471-669A-48	Sequence 48, Appl
C 39	41	3.7	1065	4	US-09-976-594-833	Sequence 833, App
40	41	3.7	4078	4	US-09-016-434-1120	Sequence 1120, Ap
41	41	3.7	5926	4	US-09-917-254-41	Sequence 41, Appl
C 42	40.8	3.7	843	4	US-09-724-797-33	Sequence 33, Appl
43	40.8	3.7	1023	4	US-09-902-540-3025	Sequence 3025, Ap
44	40.8	3.7	3073	4	US-09-902-540-643	Sequence 643, App
45	40.4	3.7	486	4	US-09-252-991A-4065	Sequence 4065, Ap

ALIGNMENTS

RESULT 1

US-09-902-540-5939

; Sequence 5939, Application US/09902540

; Patent No. 6833447

; GENERAL INFORMATION:

; APPLICANT: Goldman, Barry S.

; APPLICANT: Hinkle, Gregory J.

; APPLICANT: Slater, Steven C.

; APPLICANT: Wiegand, Roger C.

; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof

; FILE REFERENCE: 38-10(15849)B

; CURRENT APPLICATION NUMBER: US/09/902.540

; CURRENT FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/217,883

; PRIOR FILING DATE: 2000-07-10

; NUMBER OF SEQ ID NOS: 16825

; SEQ ID NO 5939

; LENGTH: 1350

; TYPE: DNA

; ORGANISM: Myxococcus xanthus

US-09-902-540-5939

Query Match 4.2%; Score 45.8; DB 4; Length 1350;
Best Local Similarity 45.6%; Pred. No. 0.28;
Matches 161; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

Qy	382	GAGCGCATCCAGCCGCGGTGGCGGTTCGAGTGGTCTCCAGGGTGGCGAACAGATGCTG	441
Db	361	GAAGCGGTGCGCGCGCGGACGAGGGCGCGCGTGGCGAGGGATGCCGAGCGCGCTG	420
Qy	442	GTGATCAAGGCGTTTCACGGCAGAGACCCAGCGCGAGTACCACCTGTGGCTGTGCTCAGGCG	501
Db	421	TCCACCATCTCAAGGGGTGAGACCTGGGAGCAACGTGGCGAGGTGTGCGCGCGCC	480
Qy	502	TCCATCTGCCCTGGCGTGCACCTTCCAGTGGCACTGAGCCCCAGCCCTCTGTGGATGAG	561
Db	481	TCCAGCGAGCAGGTGCGAGGCCACCGCGCTGGTCCAGGGCACCGCGAAGGTGAGCGAA	540
Qy	562	CCCCCTAGCCAGCTTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	621
Db	541	CAGGAGCGGCGCATCGCCACATCCCGCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG	600
Qy	622	GAGCGCGTGGCTCGGAAACGTGCGATGAGCGGTGAGCGCCAGCTGGACTCCATTAGCGAG	681
Db	601	CAGGCGCGCGGAGATGCGCGCGGATGTCGGCGGAGAGATGCGAGGAGGAGGAGGAGGAGG	660
Qy	682	CTGCAGAGGAGGAGCGCGCTCGCAGCGCTCGCAGCGCTCGCAGGAGGAGGAGGAGGAGG	734
Db	661	CGCGCGCGCTGCGCGACGTGGTGGTCTCCAAACGGGAGCTCGCGGAGGCGGT	713

RESULT 2
US-09-902-540-192/c
; Sequence 192, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 192
; LENGTH: 1351
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-192

Query Match 4.2%; Score 45.8; DB 4; Length 1351;
Best Local Similarity 45.6%; Pred. No. 0.28;
Matches 161; Conservative 0; Mismatches 192; Indels 0; Gaps 0;
QY 382 GAGCGCATCCAGCCCGGTGGCGGTTCCGAGGTGCTCCAGGTCGCGCAACAGATGCTG 441
DB 991 GAAGCGGTGCGGCGCGGACAGAGGCGCGGCTGGGAGGATGCCGAGCGCGCTG 932
QY 442 GTGATCAAGGGCTTCACGCGCAAGAACAGAGGCGGAGTACCACTGTGGCTGCTCAGGGC 501
DB 931 TCCACCATCTCAAGGGGTTGAGGACCTTGGGACCACTGCGCGAGGTGTGCGGCGCC 872
QY 502 TCATCTGCTGCGGCTGCCACCTTCAGGTGGCACTGAGCCAGCCCTCTGTGATGAG 561
DB 871 TCCAGCGAGCAGTTCAGGCCACGCGGCTGCTCCAGGACCGCGCAAGGTGAGCGAA 812
QY 562 GCCCTCAGCCAGCTTGGCCCGGAGGAGCGCCAGGAGGCTGACCTGCACCTACTGTGG 621
DB 811 CAGGAGCGGGCCATCGCCACATCCCGCGCCGAGCAGGCGCGCCAGCGCGCTGGCG 752
QY 622 GAGGCGCTGGCTCGAAACGTGCGATGAGCGGTGAGCCCGCTGCGACTCCATTAGCGAG 681
DB 751 CAGGCGCGCGGAGATGCGCGCATGTCGCGGACAGATGCGAGGCCACCTCCGAGCAG 692
QY 682 CTGCGAGGAGGAGCGCGCTCGCAGCGCTGCCACAGGAGGCGAGGAGGT 734
DB 691 GCGCGCGCGCTCGCGACGTGTGCGCTCCAAACGGGCGAGCTCGCGGAGGCGGT 639

RESULT 3
US-09-902-540-3277/c
; Sequence 3277, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 3277
; LENGTH: 2112
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-3277

Query Match 4.1%; Score 45; DB 4; Length 2112;
Best Local Similarity 46.7%; Pred. No. 0.48;
Matches 178; Conservative 0; Mismatches 200; Indels 3; Gaps 1;
QY 217 GAACCGGTTACCTGCCCCCAGACAGTGGCGGCTTCAGAGCCACCGAGCCTGTGCTCCC 276
DB 1494 GAAGCGGTAAACCCCGCCCCCAGCAGCCGCGATGATGGCGATGACCGCGCGCGCAGGGC 1435
QY 277 CAGCCTCAGCCCTGAGAGCCGCGAGGTGGCAGCTGGTGAAGATGTCTCTTGAGGCTT 336
DB 1434 GAACGGGTCCCGCCCGCCGTTAGCGGCTTCCCGATGGCGAAGATGCCGACAGCGAGTC 1375
QY 337 GAGGTGGTGGCTGAGGCTGGTGGGTATCATCTGGGCACAAAGGAATGAGCGCATCCAGCCC 396
DB 1374 GAGCCTGTCCGCGATACCGATGAGCGCGCCGCGGTCTTGGCTGGGCGACGCGCTCTCCGGC 1315
QY 397 GGTGGCGGTTCCGAGGTGGTCTCCAGGCTCGGCAACAGATGCTGGTGATCAAGGGCTTC 456
DB 1314 GCCAGCGGCGAGGTAGTCTCCGCGATGGCCAGAGCCACGGCGTCCGGCTACCGCCGC 1255
QY 457 ACGGCAGAAAGACAGGGCGAGTACCACTGTGGCTGGCTCAGGGCTCAATCTGCCCTGGC 516
DB 1254 GCGGGCGTACTCCCGGCCCATGATGCCCTGGAGCTCGGGGAACTCGGCCA---CCATGCC 1198
QY 517 GCTGCCACCTTCCAGGTGGCACTGAGCCCGCCTCTGTGGATGAGGCCCTTCAGCCCGCAGC 576
DB 1197 GGTGACGAGGTGCGGCTTGGCGAGGTGGCGCGCGCTCGATGATAGCGGCTCTCCCTGC 1138
QY 577 TTGCCCCCGAGGCGAGCCAG 597
DB 1137 TCGCCCCGCTCTCTGGGCCAG 1117

RESULT 4
US-09-902-540-1151
; Sequence 1151, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 1151
; LENGTH: 17480
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-1151

Query Match 4.1%; Score 45; DB 4; Length 17480;
Best Local Similarity 46.7%; Pred. No. 0.79;
Matches 178; Conservative 0; Mismatches 200; Indels 3; Gaps 1;
QY 217 GAACCGGTTACCTGCCCCCAGACAGTGGCGGCTTCAGAGCCACCGAGCCTGTGCTCCC 276
DB 7398 GAAGCGGTAAACCCCGCCCCCAGCAGCCGCGATGATGGCGATGACCGCGCGCGCAGGGC 7457
QY 277 CAGCCTCAGCCCTGAGAGCCCGCGAGGTGGAGCTGGTGAAGATGTCTCTTGAGGCTT 336
DB 7458 GAACGGGTCCGCGCGCCCGCTTGGCGGCTTGGCGATGGCGAAGATGCCGACAGCGAGTC 7517
QY 337 GAGGTGGTGGCTGAGGCTGGTGGGTATCATCTGCGCAAGGAATGAGCGCATCCAGCCC 396
DB 7518 GAGCCTGTCCGCGATACCGATGAGCGCGCCGCGGTCTTGGCTGGGCGAGCGCGTCTCCGGC 7577
QY 397 GGTGGCGGTTCCGAGGTGGTCTCCAGGCTCGGCAACAGATGCTGGTGATCAAGGGCTTC 456

Db 7578 GCCAGCGGCGAGGTAGTCTCCGATGCGCAGACCGCGGTCCGGCTTACCCGCC 7637
Qy 457 ACGGAGAGACACAGGCGAGTACCACTGTGCGCTTGGCTCAGGGGTCCATCTGCCCTCGG 516
Db 7638 GCGGCGGTACTCCGCGCCATGATGCCCTGGAGCTCGGGAACTCGCCCA---CCATGCC 7694
Qy 517 GCTGCACCTTCCAGGTGGCACTGAGCCAGCCTCTGTGGATGAGGCCCCCTCAGGCCAGC 576
Db 7695 GGTGACGAGGTGCGGCTTGGCCAGAGGTGGCGCGCGCTCGATGGTAGCGCGCTCCCGCTGC 7754
Qy 577 TTGCCCCCGGAGGACGCCAG 597
Db 7755 TCGCCCCCGCTCTCGGCCAG 7775

RESULT 5

US-09-144-085-3
; Sequence 3, Application US/09144085
; Patent No. 628099
; GENERAL INFORMATION:
; APPLICANT: Gustafsson, Claes
; APPLICANT: Betlach, Mary C.
; APPLICANT: Ahlvey, Gary
; APPLICANT: Julien, Bryan
; APPLICANT: Ziermann, Rainer
; TITLE OF INVENTION: SORANGIUM POLYKETIDE SYNTHASES AND ENCODING DNA
; FILE REFERENCE: 30062-20020.20
; CURRENT APPLICATION NUMBER: US/09/144,085
; EARLIER FILING DATE: 1998-08-31
; EARLIER APPLICATION NUMBER: 09/010,809
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 33529
; TYPE: DNA
; ORGANISM: Sorangium cellulosum
US-09-144-085-3

Query Match 4.1%; Score 45; DB 3; Length 33529;
Best Local Similarity 44.8%; Pred. No. 0.92;
Matches 298; Conservative 0; Mismatches 360; Indels 7; Gaps 3;
Qy 74 CGGTGGAGGTGAAGAGGGGGCCACAGGCGCTGTGCGGCTGTGCCAGAGCTGTGCGCTG 133
Db 13643 GCGCGTGGCGCGCGCCAGACGACGCGCTCGAGGACTTGGCGCGCGCGCTGTGG 13702
Qy 134 GACCGAGTGTGTGTGATGGCTTGGCGCCCGCGGGAGACCTACCGCTTCCGTGTGGCAG 193
Db 13703 GCTTGTCCGCGCGCGGCGAAGCGAGACCCCGAAGCGGGGCTGCGCTTGTGATGTGG 13762
Qy 194 CTGTGGCGCTGTGGGTGTGGGAACCGTTTCACTGCCCGCAGACAGTGGCGCTTGAG 253
Db 13763 GGACGAGCCCGTGAGCGCTGGGCTGTGCGCGGCGGCTGGCGAGCGCGCGAGCCGG 13822
Qy 254 AGCCACCGAAGCTGTGCTCTCCAGCCCTCAGCCCTCAGAGCGCGGAGGTGGCAGCTG 313
Db 13823 AGCTTGGCTGCGCGGGGCGCTGCGCTGGCGCGCGCTGTGTGCGCAGAGCGGTAG 13882
Qy 314 GTGAAGATGTCTCTCTGAGCTTGAAGTGTGGCTGTGAGCTGTGAGTCACTCTGGCA 373
Db 13883 CGGAAGAGCTCACCGAGCCCGAGCTGGACCTTGGCGGCGACGCTGTGTGTGACCGGG 13942
Qy 374 AGGAATGAGCGCATCCAGCCGCTGGCGGTTCAGTGTGTCTCCAGGGTCCGCAAC 433
Db 13943 GGACAGGGAGCTGGGTTCAGGCGGTTCGCGGCGCACTGTGTGCGCGCAGCGGGTGC 14002
Qy 434 AGATCTGTGTA--TCAAGGGCTTCAAGGAGAGACAGGCGGAGTACCACTGTGGCT 491
Db 14003 ACCTTGCTGACGTGCGCGCGGGCTGGAGGCGCGCGGGCGCGAGCTTGTGGAT 14062

Qy 492 GGCTCAGGGTCCATCTGCGCTGCCACCTTCCAGGTGGGACTGAGCCCGACCTC 551
Db 14063 CGCTCGCGAGCTCGCGCCGAGACGGTGAACGGTGGCGCGCTGCAGCTGCAGAGC--G 14120
Qy 552 TGTGATGAGGCGCCCTCAGCCAGCTTGCCTCCCGAGGAGCCAGGAGGGTGAACCTGCA 611
Db 14121 GGAGAGGTTCGCGCTGTGCTGGCGGATGACGCGGCGCGCTGAGCGCGGTGT 14180
Qy 612 CTAAGTGGAGGCGCTTGGCTCGGAAACGTGCAATGAGCCGTGAGCCACCTGAGATC 671
Db 14181 GCACCTCGCGCGCTCGACGCGGTGTCTCGCGGCGAGAGCGCGAGCGCTCTC 14240
Qy 672 CATTAAGCAGTGCAGAGGAGGCGCGCTCGCAGCGCTGCGCAGGAGGAGAGGA 731
Db 14241 --GCGGTGTCTGGCGCGAAGGTGGACGGGCGCTGCACCTGACAGCTGACCGGGA 14297
Qy 732 GGTGG 736
Db 14298 GGTGG 14302

RESULT 6

US-09-252-991A-15050/c
; Sequence 15050, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15050
; LENGTH: 765
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15050

Query Match 4.1%; Score 44.8; DB 4; Length 765;
Best Local Similarity 44.0%; Pred. No. 0.42;
Matches 190; Conservative 0; Mismatches 242; Indels 0; Gaps 0;
Qy 46 GATGAGCGGTGTCTCTGTGGCTACCGCTGGAGGTGAAGAGGGGGCCACAGCCAG 105
Db 641 GCTGTGCGGTGTGGGTGAGGGCGACCGCGCGAGTTGACAGGGCGCTTGTGCGCTT 582
Qy 106 TGGCGGTGTGCCACGAGCTGTGCTGGACCCGAGTGTGTGTGGATGGCTGGCCCCC 165
Db 591 GCCCGGCTTCCAGCGCGTAGCCGGGGTGGCGGTTTGGCAGAGACGACCTCGCGT 522
Qy 166 GGGAGACCTACCGCTTCCGTGTGGCAGCTGTGGGCCCTGTGGGTGTGTGGGAAACGGTT 225
Db 521 CTGCGCCCGAGCGGTGGCCATCCCTTCGCGGATTCGTAGCGAGGCTGGCGAGGCCCTG 462
Qy 226 CACCTGCCCGCAGACAGTGGCTTGCAGAGCCAGAGCTGTGCTCCCGAGCCCTCA 285
Db 461 GAGCAGCGCTTGGCGGCGGCGAGGGCGCTGTGTGGAGGTGGCGGTGAGGATGGC 402
Qy 286 GCCCTGAGAGCGCGCAGGTGGCAGTGTGTAAGATGTCTCTCTGGAGCTTGAAGTGGT 345
Db 401 GTTGATCGCTCGGTGAGTTGCGGAGTTGTCGGATCTCTGTGCGGCTTGCAGCGGC 342
Qy 346 GCTGAGGCTGTGAGGTCACTGCGACAAGGGAATGGAGCGCATCCAGCCCGGTGGCGG 405
Db 341 GCGCGCTGGAGAGTTGCGGTGTTCGGCGCGCGAGGGCGGCGACTTCGCTGGCGAG 282
Qy 406 TTCAGGTGTCTCCAGGTTCGCAACAGATGTGTGTGATCAAGGGCTTACGGCAGNA 465

Db 281 GTCTTGACGTTTCACGCGCTGGCGTTGCTGCTGGTGTTCGCGGGCGGCGCTGGGCGAG 222
QY 466 GACCAGGGCGAG 477
Db 221 GTCGAGGACCAG 210

RESULT 7
US-09-252-991A-14932/c
; Sequence 14932, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14932
; LENGTH: 1329
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14932

Query Match 4.1%; Score 44.8; DB 4; Length 1329;
Best Local Similarity 44.0%; Pred. No. 0.47;
Matches 190; Conservative 0; Mismatches 242; Indels 0; Gaps 0;
QY 46 GATGAGGCGGTGCTCTGTGGCTACCGCTGAGGCTGAAGGAGGGGGCCACAGGCCAG 105
Db 1146 GCTGTGCGGTTGGGTTGAGGGCGACCGCGGCGAGTTGCGAGCGCGCGTTGTGCGGCTT 1087
QY 106 TGGCGGCTGTGCCACGAGCTGTGCTCGACCGAGTGTGTGGTGTGATGGCTTGCGCCCGC 165
Db 1086 GCGCGCGACTTCCAGGCGGTAGCCGGGGGTGGCGGGTTTGGCAGGACGACCTTGGCGGT 1027
QY 166 GGGGAGACCTACCGCTTCCGTGTGGCAGCTGTGGGCCCTGTGGGTCTTGGGGAACCGGTT 225
Db 1026 CTGCGCCAGCGGTGGCCATGCTTCGCGACCTTCGTAGCCGAGGCTTGGGAGGCCCTG 967
QY 226 CACCTGCCCCAGACAGTGGCTTGCAGAGCCACGAGCCTGTGCCCTCCCGACCCCTCA 285
Db 966 GAGCAGGCTTGGCGCGCGGCGAGGCGGCTGTGCTGGAGGTGGGCGGTGAGATGGC 907
QY 286 GCGGCTGAGAGCGGCGAGGTGGCAGCTGTGGAAGATGTCTCTTGGAGCTTGGGTGGTG 345
Db 906 GTTGATCGCTCGGTGAGTTGCGCGAGTTGCTGCGGATCTCTGTGCGGCTGGCAGGCGG 847
QY 346 GCTGAGGCTGTGAGGTATCTGGCACAAGGATGAGGCGCATCCAGCCCGTGGGGG 405
Db 846 GCGCGCTGAGAGAGTTCCGCGGTGTTCGCGCGCGCGAGGCGCGACTTCGCTGGCGAG 787
QY 406 TTCGAGGTGGTCTCCAGGGTGGGCAACAGATGCTGGGTGATCAAGGGCTTCAACGCGAGAA 465
Db 786 GTCTTGACAGTTTCCAGGCGCTGGCGTTGCTGTGGTGTTCGCGGGCGGCTGGGCGAG 727
QY 466 GACCAGGCGAG 477
Db 726 GTCGAGGACCAG 715

RESULT 8
US-09-252-991A-15189/c
; Sequence 15189, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 15189
; LENGTH: 5121
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-15189

Query Match 4.1%; Score 44.8; DB 4; Length 5121;
Best Local Similarity 44.0%; Pred. No. 0.66;
Matches 190; Conservative 0; Mismatches 242; Indels 0; Gaps 0;
QY 46 GATGAGGCGGTGCTCTGTGGCTACCGCTGAGGCTGAAGGAGGGGGCCACAGGCCAG 105
Db 4843 GCTGTGCGGTTGGCGTTGAGGGCGACCGCGCGAGTTGAGGCGGCGGTTGTGCGGCTT 4784
QY 106 TGGCGGCTGTGCCACGAGCTGTGCTCGACCGAGTGTGTGGTGTGATGGCTTGGCCCGC 165
Db 4783 GCGCGCGACTTCCAGGCGGTAGCCGGGGTGGCGGGTTTGGCAGGAGCAGCCCTGCGGT 4724
QY 166 GGGGAGACCTACCGCTTCCGTGTGGCAGCTGTGGGCCCTGTGGGTCTTGGGGAACCGGTT 225
Db 4723 CTGCGCCAGCGGTGGCCATGCTTCGCGGACTTCGTAGCCGAGGCTTGGCGAGGCCCTG 4664
QY 226 CACCTGCCCCAGACAGTGGCGCTTGCAGAGCCACGAGCCTGTGCCCTCCCGACCCCTCA 285
Db 4663 GAGCAGGCTTGGCGGCGGCGAGGCGGCTGTGCTGGAGGTGGGCGGTGAGATGGC 4604
QY 286 GCGGCTGAGAGCGGCGAGGTGGCAGCTGTGTGAAGATGTCTCTTGGAGCTTGGAGGTGTG 345
Db 4603 GTTGATCGCTCGGTGAGTTGCGCGAGTTGCTGCGGATCTCTGTGCGGCTGGCAGGCGG 4544
QY 346 GCTGAGGCTGTGAGGTATCTGGCACAAGGATGAGGCGCATCCAGCCCGTGGGCGG 405
Db 4543 GCGCGCTTGGAGCAGTTTGGCGGTGTTTGGCGGCGCGAGGCGGCGACTTCGCTGGCGAG 4484
QY 406 TTCGAGGTGTCTCCAGGGTGGGCAACAGATGCTGTGATCAAGGGCTTCAAGGCGAGAA 465
Db 4483 GTCTTGACAGTTTCCAGGCGCTGGCGTTGCTGTGGTGTTCGCGGGCGGCTGGGCGAG 4424
QY 466 GACCAGGCGAG 477
Db 4423 GTCGAGGACCAG 4412

RESULT 9
US-09-252-991A-14239
; Sequence 14239, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14239
; LENGTH: 735
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14239

Qy	544	CAAGCCTCTGTGGATAGAGCCCTCTAGCCACGCTTGCCCTCCGAGGCGACGCCAGGAGGGT	603
Db	220	CATGGCGGCGATCAGCGGCTAGCGGCTTCATCTCCGCCACGCTCCAGCGCAGGCGGTT	161
Qy	604	GACCTGCACCTAC	616
Db	160	GGCGGCTTCCTTC	148

```

RESULT 13
US-09-252-991A-14994
; Sequence 14994, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14994
; LENGTH: 2502
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14994

```

Query Match	3.9%	Score 42.6	DB 4	Length 2502
Best Local Similarity	51.3%	Pred. No. 1.8		
Matches 99	Conservative 0	Mismatches 94	Indels 0	Gaps 0
QY	424	GGTCGGCAACAGATGCTGGTGATCAAGGGCTTCACGCGCAGAGACACGAGCGCGAGTACCAC	483	
Db	694	GCTCAGCTGCGCGCGGCTCGTTCGGGGCTTTCCGACGAAACAGCGGTCGGTCAGCGG	753	
QY	484	TGTGGCCTGGGCTCAGGGGCTCCATGTCCTTCGCGCTGCCACCTTCAGGTGGCACTGAGC	543	
Db	754	CGTGCACGCTCCTCGGGCGTTCCTTACCTTTGCGGTTGAAGCCAGCCAGGTTGCGCACACC	813	
QY	544	CCAGGCTCTGTGGATAGGGCCCTCCAGCCAGCTTCGCCCGGAGGACGCCAGAGGGT	603	
Db	814	CATGCCGCCATCAGGCGGTAGCGCCGCTCATCTCCGCCACGCTCCAGCGAGGGCGTT	873	
QY	604	GACCTGCACCTAC	616	
Db	874	GCGGCTTCTTC	886	

```

RESULT 14
US-09-252-991A-14750/c
; Sequence 14750, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14750
; LENGTH: 2505
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14750

```

Query Match	3.9%	Score 42.6;	DB 4;	Length 2505;
Best Local Similarity	51.3%;	Pred. No. 1.8;		
Matches 99;	Conservative 0;	Mismatches 94;	Indels 0;	Gaps 0;
Qy	424	GGTCGGCAACAGATGCTGGTGATCAAGGGCTTCACGGCAGAGACACGAGCGGAGTAGTACCAC	483	
Db	1824	GCTCAGGTGGGGGGTCTGTCGTCGGGCTTTCGCACGGACAGCGGGTCGGTCAGCGG	1765	
Qy	484	TGTGGGCTGCTCAGGGCTCCATCTGCCTCGGGCTGCGCACTTCCAGGTGGCACTGAGC	543	
Db	1764	CGTGCCAGCCTCTCGGCGTCTCTCACTTGCGTTGAAGCCAGCCAGGTTCGCACACC	1705	
Qy	544	CCAGCTCTGTGATCAGGCCCTTCAGCCACAGCTTGCCTCCCGAGGACGCCAGGAGGCT	603	
Db	1704	CATGGCGGCATCAGCGGTGAGCGCCGCTCATCTCCGCCACGCTCCAGCGGCGCGTT	1645	
Qy	604	GACCTGCACTAC	616	
Db	1644	GGCGGCTTCCTTC	1632	

RESULT 15
US-09-949-016-2188
; Sequence 2188, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CU001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 2188
; LENGTH: 3397
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-2188

Query Match	3.8%;	Score 42;	DB 4;	Length 3397;
Best Local Similarity	51.7%;	Pred. No. 2.7;		
Matches 121;	Conservative 0;	Mismatches 110;	Indels 3;	Gaps 1;
QY	297	CGGCAGAGTGGCAGCTGGTGAAGATGTCTCTCGAGCTTGAGGTGGTGGCTGAGGCTGG	356	
Db	888	CTGATGGTGGCGCAAGGACCGCGGTTCAAATGTGAGGTCTCAGATGAGAAATGT	947	
QY	357	TGAGGTTCATCTGGCAACAGGGGAATCGAGCGCATCCAGCCGGTGGCGGTTGAGGTGGT	416	
Db	948	TCGGGGTGTGTGGCTGAAGAAATGGGAAGAGCTGTTGCCCGCAGCCGCATAAAGGTGTC	1007	
QY	417	CTTCCAGAGTGGGCAACAGATCTCTGGTGATCAAGGGCTTCAGGCAGAGAAGACCGGGCGA	476	
Db	1008	CACATCTGGCGGGTCCAAACTGACCATTCACACGTCTCACCTGCCGAGAGGCTGA	1067	
QY	477	GTACCACTGTGGCCTGGTCTCAGGGCTTCATCTGGCCCTCGGCTGCCACTTCCA	530	
Db	1068	CTACAGCTTTG---TGCOCGAGGGCTTCGCTTGAACCTGTGAGCAAGCTCCA	1118	

Search completed: March 21, 2005, 14:39:32
Job time : 207.228 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:32:32 ; Search time 627.052 Seconds
(without alignment)
9503.485 Million cell updates/sec

Title: US-10-077-130-6_COPY_2000_3000

Perfect score: 1001

Sequence: 1 CGCTCACCATCCGGAGGTG.....CTGAGGCGAGGGGCAATGCC 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	23907	13 US-10-077-130-6	Sequence 6, Appli
2	1001	100.0	24120	13 US-10-077-130-4	Sequence 4, Appli
3	983	98.2	20489	18 US-10-723-860-5701	Sequence 5701, Ap
4	830.8	83.0	2768	17 US-10-120-988-426	Sequence 426, App
5	191	19.1	578	13 US-10-027-632-272054	Sequence 272054,
6	191	19.1	578	17 US-10-027-632-272054	Sequence 22, Appl
7	88.4	8.8	1645	9 US-09-726-643-22	Sequence 22, Appl
8	88.4	8.8	1645	13 US-10-042-141-22	Sequence 22, Appl
9	88.4	8.8	1645	19 US-10-919-272-22	Sequence 22, Appl
10	88.4	8.8	2170	17 US-10-264-049-850	Sequence 850, App
11	86.2	8.6	2534	17 US-10-108-260A-1946	Sequence 1946, Ap

12	86.2	8.6	7564	18 US-10-476-397-15	Sequence 15, Appl
13	78.6	7.9	14061	17 US-10-093-463-73	Sequence 73, Appl
14	78.6	7.9	14109	17 US-10-093-463-71	Sequence 71, Appl
15	67.8	6.8	2254	9 US-09-726-643-44	Sequence 44, Appl
16	67.8	6.8	2254	13 US-10-042-141-44	Sequence 44, Appl
17	67.8	6.8	2254	19 US-10-919-272-44	Sequence 44, Appl
18	64.8	6.5	716	10 US-09-822-846-117	Sequence 117, App
19	63.4	6.3	20565	10 US-09-764-891-6045	Sequence 6045, Ap
20	63.4	6.3	20565	14 US-10-091-438-270	Sequence 270, App
21	54	5.4	1266	10 US-09-822-846-116	Sequence 116, App
22	53.4	5.3	2884	17 US-10-094-749-1001	Sequence 1001, Ap
23	52.8	5.3	634	13 US-10-027-632-27250	Sequence 27250, A
24	52.8	5.3	634	17 US-10-027-632-27250	Sequence 27250, A
25	52.8	5.3	2695	15 US-10-037-270-473	Sequence 473, App
26	52.8	5.3	2695	17 US-10-117-722-473	Sequence 473, App
27	52.6	5.3	400	17 US-10-242-535A-8373	Sequence 8373, Ap
28	52.6	5.3	400	17 US-10-085-783A-8373	Sequence 8373, Ap
29	52.2	5.2	4953	18 US-10-483-506-29	Sequence 29, Appl
30	52	5.2	1512	18 US-10-437-963-17331	Sequence 17331, A
31	52	5.2	1640	17 US-10-425-114-26270	Sequence 26270, A
32	52	5.2	1659	18 US-10-437-963-17332	Sequence 17332, A
33	51.8	5.2	2076	18 US-10-357-930-22530	Sequence 22530, A
34	51.8	5.2	2076	18 US-10-357-930-28365	Sequence 28365, A
35	51.8	5.2	3408	14 US-10-198-846-11405	Sequence 11405, A
36	51.6	5.2	3575	17 US-10-641-643-1395	Sequence 1395, Ap
37	51.6	5.2	3575	18 US-10-723-860-1253	Sequence 1253, Ap
38	51.6	5.2	3597	13 US-10-098-841-84	Sequence 84, Appl
39	51.6	5.2	3609	18 US-10-723-860-5754	Sequence 5754, Ap
40	51.6	5.2	9591	10 US-09-764-891-6043	Sequence 6043, Ap
41	51.6	5.2	9591	14 US-10-091-438-268	Sequence 268, App
42	51.6	5.2	12415	10 US-09-764-891-6044	Sequence 6044, Ap
43	51.6	5.2	12415	14 US-10-091-438-269	Sequence 269, App
44	50.4	5.0	1839	11 US-09-758-759-150	Sequence 150, App
45	50.4	5.0	3833	9 US-09-969-708-391	Sequence 391, App

ALIGNMENTS

RESULT 1

US-10-077-130-6
; Sequence 6, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapellner-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047PARCP1(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 23907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-077-130-6

Query Match 100.0%; Score 1001; DB 13; Length 23907;
Best Local Similarity 100.0%; Pred. No. 8.8e-248;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 CGCTCACCATCCGGAGGTGCGCCGACCTGCGACGGCGCGAGTGAAGTTCGTGGCCA 60
Db 2000 CGCTCACCATCCGGAGGTGCGCCGACCTGCGACGGCGCGAGTGAAGTTCGTGGCCA 2059
Oy 61 ACGCCATTGAGCAGCATCCGATGAGGTCCCGCGCGCGCGCGCTGACTGCCAACA 120
Db 2060 ACGCCATTGAGCAGCATCCGATGAGGTCCCGCGCGCGCGCGCTGACTGCCAACA 2119

Qy	121	AGCGCCAGCCGACGCTCCCGGAGAGGTGCTGGCTCGGCTGCACGAGAGGCGCAGCTGC	180
Db	2120	AGCGCCAGCCGACGCTCCCGGAGAGGTGCTGGCTCGGCTGCACGAGAGGCGCAGCTGC	2179
Qy	181	TGGCTGAGCTGTCAGATCAGGCTGCGGCTGTGACGTGGCTGAAGAGATGGTCGCACTGT	240
Db	2180	TGGCTGAGCTGTCAGATCAGGCTGCGGCTGTGAAGTGGCTGAAGATGGTCGCACTGT	2239
Qy	241	CCCAGGCCCCAAGTATGAGGTGAGGCATCGGCGCGGCGCGGCTGCTCCTTGTGCGAG	300
Db	2240	CCCAGGCCCCAAGTATGAGGTGAGGCATCGGCGCGGCGCGGCTGCTCCTTGTGCGAG	2299
Qy	301	ATGTGGCCCGGAGGATGACGSCCTCTACGATGCGTACGCGCGGGGGCGCATCGCCT	360
Db	2300	ATGTGGCCCGGAGATGATCGAGCCTCTACGATGCGTACGCGCGGGGGCGCATCGCCT	2359
Qy	361	ACCAGCTCTCGTGCAGAGCCTCGCGGCTTTCTGCACAAGGACATCGCGGGGAGCTGTG	420
Db	2360	ACCAGCTCTCGTGCAGAGCCTCGCGGCTTTCTGCACAAGGACATCGCGGGGAGCTGTG	2419
Qy	421	TGGATGCGGTGGCTGGGGGCCCGGCGCAGTTTGATGTGTAGA CTTCCGAAGCCACGTC	480
Db	2420	TGGATGCGGTGGCTGGGGGCCCGGCGCAGTTTGATGTGTAGA CTTCCGAAGCCACGTC	2479
Qy	481	ACGTGCATCTGTACAAAGATGGCATGGAGCTGGGCCACTCGGTGAGCGCTTCTTTCAGG	540
Db	2480	ACGTGCATCTGTACAAAGATGGCATGGAGCTGGGCCACTCGGTGAGCGCTTCTTTCAGG	2539
Qy	541	AGGATGTGGGACCGCGGACCGGCTGGTGGGAGGCA CAGTCA CAGGACAGGATGAAGGCA	600
Db	2540	AGGATGTGGGACCGCGGACCGGCTGGTGGGAGGCA CAGTCA CAGGACAGGATGAAGGCA	2599
Qy	601	CCTACTCTGCGGCTGGGCGAGACTCTGTGGAAGTTCGCGCTCCGCGTCTCTGAGCCCA	660
Db	2600	CCTACTCTGCGGCTGGGCGAGACTCTGTGGAAGTTCGCGCTCCGCGTCTCTGAGCCCA	2659
Qy	661	AGTGTGTTTGTGTAAGGACAGCTGSCACCGCAGGAAGCTGCGAGGCAAGGCGAGGACCA	720
Db	2660	AGTGTGTTTGTGTAAGGACAGCTGSCACCGCAGGAAGCTGCGAGGCAAGGCGAGGACCA	2719
Qy	721	GTGCCACACTGAGCTGCGAGTGGCCAGCGCCAGACGGAGGTGACGTGTGTAACGATG	780
Db	2720	GTGCCACACTGAGCTGCGAGTGGCCAGCGCCAGACGGAGGTGACGTGTGTAACGATG	2779
Qy	781	GGAAGAGCTGAGCTCCAGCTCGAAAGTGTGCATGAGGCGCCACAGGCTGCACGCGCAGGC	840
Db	2780	GGAAGAGCTGAGCTCCAGCTCGAAAGTGTGCATGAGGCGCCACAGGCTGCACGCGCAGGC	2839
Qy	841	TGGTTGTGACGAGGCGAGGCTGCGGAGTACGCGGGAGTATAGCTGCGAGGCTGGGGGCC	900
Db	2840	TGGTTGTGACGAGGCGAGGCTGCGGAGTACGCGGGAGTATAGCTGCGAGGCTGGGGGCC	2899
Qy	901	AGGGCTCTCTTCCATCTGATGTCAAAGAGCCCAAGGTGTGTTTGGCCAAGGACAGG	960
Db	2900	AGGGCTCTCTTCCATCTGATGTCAAAGAGCCCAAGGTGTGTTTGGCCAAGGACAGG	2959
Qy	961	TGGCACACAGTGAAGTGCAGGCTGAGCGAGGGGCCAATGCC	1001
Db	2960	TGGCACACAGTGAAGTGCAGGCTGAGCGAGGGGCCAATGCC	3000

RESULT 2

US-10-077-130-4

US-10-077-130-4 : Sequence 4. Application US/10077130

Publication NO. US20020168743A1
; sequence 4, Application US/1007

; Publication No. US20
; GENERAL INFORMATION:

APPLICANT: Kanelker-Libermann, Rosanna

; APPLICANT: Kapellef=Liberman
; APPLICANT: Action Susan I

APPLICANT: Acton, Susan L.
TITLE OF INVENTION: 50070 and 10500 product mixtures

59079 and 12599, Protein K

;	TITLE OF INVENTION: Members and Use
:	FILE REFERENCE: MD12001-01755-0001 (M)

```

; FILE REFERENCE: MPI2001-047P1RCPI(M)
; CURRENT ADDICTION NUMBER: 10/10/07 130

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84


```
Qy 781 GGAAGAAGCTGAGCTCCAGCTCGAAAGTGTGCATGGAGGCCACAGGCTGCACGGCAGGC 840
Db 2851 GGAAGAAGCTGAGCTCCAGCTCGAAAGTGTGCATGGAGGCCACAGGCTGCACGGCAGGC 2910
Qy 841 TGGTTGTGACAGCAGCAGGCCAGCGGATGCCGGGGAGTATAGCTGCGAGGCTGGGGGCC 900
Db 2911 TGGTTGTGACAGCAGCAGGCCAGCGGATGCCGGGGAGTATAGCTGCGAGGCTGGGGGCC 2970
Qy 901 AGCGGCTCTCTTCATCTGATGTCAAGAGCCCAAGGTGGTGTTCGCAAGACACAGG 960
Db 2971 AGCGGCTCTCTTCATCTGATGTCAAGAGCCCAAGGTGGTGTTCGCAAGACACAGG 3030
Qy 961 TGGCACACAGTGAGTGTGAGCTGAGGCGAGGCGCAATGCC 1001
Db 3031 TGGCACACAGTGAGTGTGAGCTGAGGCGAGGCGCAATGCC 3071

RESULT 3
US-10-723-860-5701
; Sequence 5701, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; TITLE OF INVENTION: Methods for Screening for Soft Tissue Sarcoma Modulators
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5701
; LENGTH: 20489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (565)..(584)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2938)..(2955)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3329)..(3364)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4556)..(4586)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4721)..(4735)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4997)..(5011)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (5108)..(5138)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (6237)..(6300)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (7520)..(7566)
```

```
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12014)..(12037)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (12123)..(12146)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13258)..(13277)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (13918)..(13942)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15219)..(15248)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15262)..(15280)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15309)..(15381)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (15662)..(15682)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (17286)..(17303)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (17413)..(17430)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-5701

Query Match 98.2%; Score 983; DB 18; Length 20489;
Best Local Similarity 98.2%; Pred. No. 3.8e-243;
Matches 983; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

Qy 1 CGCTCACCATCCGGGAGGTGCCCGCAGCTGCACGGGGCGCAGCTGAAGTTCTGTGGCCA 60
Db 2071 CGCTCACCATCCGGGAGGTGCCCGCAGCTGCACGGGGCGCAGCTGAAGTTCTGTGGCCA 2130
Qy 61 ACGGCATTGAGAGCAGCATCCGGATGGAGTCCGGGGCGGCCCGGCTGACTGCCAACA 120
Db 2131 ACGGCATTGAGAGCAGCATCCGGATGGAGTCCGGGGCGGCCCGGCTGACTGCCAACA 2190
Qy 121 AGCCGCCAGCCGAGCTGCCCGGAGGTGCTGGCTCGGCTGCACGAGAGGGGCGAGCTGC 180
Db 2191 AGCCGCCAGCCGAGCTGCCCGGAGGTGCTGGCTCGGCTGCACGAGAGGGGCGAGCTGC 2250
Qy 181 TGGCTGAGCTGTGATGATCAGGCTCGGCTGTGAGCTGAGTGGCTGAGGATGGTGCACACTGT 240
Db 2251 TGGCTGAGCTGTGATGATCAGGCTCGGCTGTGAGCTGAGTGGCTGAGGATGGTGCACACTGT 2310
Qy 241 CCCGAGGCCCAAGTATGAGTGCAGGCATCCGGGGCGGCCCGGCTGCTTGTGCGAG 300
Db 2311 CCCGAGGCCCAAGTATGAGTGCAGGCATCCGGGGCGGCCCGGCTGCTTGTGCGAG 2370
Qy 301 ATGTGGCCCGGAGCAGTACAGGCTCTACAGTGGGTGAGGCGGGGGCGGCGCATCGCT 360
Db 2371 ATGTGGCCCGGAGCAGTACAGGCTCTACAGTGGGTGAGGCGGGGGCGGCGCATCGCT 2430
Qy 361 ACCAGCTCTCGTGCAGGCGCTTCGCGCTTTCTGCAAGGACATGCGGCGAGCTGTG 420
```


Db 2431 ACCAGCTCTCCGTGCAAGCCCTCCGCGCTTTCTGCACAAGACATGCGGGCAGCTGTG 2490
Qy 421 TGGATGCCGTGGCTGGGGCCCGCGCAGTTTGTAGTGTGAGACCTCCCGAAGCCACAGTCC 480
Db 2491 TGGATGCCGTGGCTGGGGCCCGCGCAGTTTGTAGTGTGAGACCTCCCGAAGCCACAGTCC 2550
Qy 481 AGCTGCATCTGTGTAAGGATGCAATGAGCTGGGCACTCCCGGTGAGCGCTTCTTGCAGG 540
Db 2551 AGCTGCATCTGTGTAAGGATGCAATGAGCTGGGCACTCCCGGTGAGCGCTTCTTGCAGG 2610
Qy 541 AGGATGTGGGACCGCGGACCGGCTGGTGGGAGCCACAGTACACAGGAGGATGAGGCA 600
Db 2611 AGGATGTGGGACCGCGGACCGGCTGGTGGGAGCCACAGTACACAGGAGGATGAGGCA 2670
Qy 601 CCTACTCTCCGCGTGGCGGAGGACTCTGTGGACTTCCCGGCTCCCGGCTCTGTAGCCCA 660
Db 2671 CCTACTCTCCGCGTGGCGGAGGACTCTGTGGACTTCCCGGCTCCCGGCTCTGTAGCCCA 2730
Qy 661 AGGTGGTGTCTTAAGGAGCAGCTGGGCAACGAGGAGCTGCAGGAGGAGGAGGCA 720
Db 2731 AGGTGGTGTCTTAAGGAGCAGCTGGGCAACGAGGAGCTGCAGGAGGAGGAGGCA 2790
Qy 721 GTGCCACACTGAGCTGCGAGGTGGCCAGGCGCCAGGAGGAGGAGGAGGAGGATG 780
Db 2791 GTGCCACACTGAGCTGCGAGGTGGCCAGGCGCCAGGAGGAGGAGGAGGAGGATG 2850
Qy 781 GGAAGAAGCTGAGCTCCAGCTCGAAGTGTGATGGAGGCCACAGGCTGCACGCGCAGGC 840
Db 2851 GGAAGAAGCTGAGCTCCAGCTCGAAGTGTGATGGAGGCCACAGGCTGCACGCGCAGGC 2910
Qy 841 TGGTGTGACAGGAGGCGGAGGAGGCGGATGCGGGAGTATAGCTGCGAGGCTCGGGGCC 900
Db 2911 TGGTGTGACAGGAGGCGGAGGAGGCGGATGCGGGAGTATAGCTGCGAGGCTCGGGGCC 2970
Qy 901 AGCGCTCTCTCTCCATCTGATGTCAAGAGCCCAAGGCTGGTGTGGCCAGGACAGG 960
Db 2971 AGCGCTCTCTCTCCATCTGATGTCAAGAGCCCAAGGCTGGTGTGGCCAGGACAGG 3030
Qy 961 TGGCACACAGTGAGGTGAGGCTGAGGCTGAGGCGAGGCGCAATGCC 1001
Db 3031 TGGCACACAGTGAGGTGAGGCTGAGGCTGAGGCGAGGCGCAATGCC 3071

RESULT 4
US-10-988-426
; Sequence 426, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyan
; APPLICANT: Wang, Dunrui
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt FL_genes Version 2.0
; SEQ ID NO 426
; LENGTH: 2768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72) .. (2738)
US-10-120-988-426

Query Match 83.0%; Score 830.8; DB 17; Length 2768;

Best Local Similarity 95.3%; Pred. No. 5.4e-204;
Matches 856; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 1 CGCTCACATCCGGAGGTGCCCCAGCTTGCACGGGGCGCAGCTGAAGTTCGTGGCCA 60
Db 1861 CGCTCACATCCGGAGGTGCCCCAGCTTGCACGGGGCGCAGCTGAAGTTCGTGGCCA 1920
Qy 61 ACGGCATTTAGAGCAGCATCCGGATGAGGTCCGGCGGGCCAGGGCTGACTGCCAACA 120
Db 1921 ACGGCATTTAGAGCAGCATCCGGATGAGGTCCGGCGGGCCAGGGCTGACTGCCAACA 1980
Qy 121 AGCCGACGCGCAGCTGCCCGGAGGTGCTGGCTCGGCTGCAACAGGAGGCGCAGCTGC 180
Db 1981 AGCCGACGCGCAGCTGCCCGGAGGTGCTGGCTCGGCTGCAACAGGAGGCGCAGCTGC 2040
Qy 181 TGGCTGAGCTGTGATGATCAGGCTGCGGCTGTGAGTGGCTGGAAGGATGCTGCACTGT 240
Db 2041 TGGCTGAGCTGTGATGATCAGGCTGCGGCTGTGAGTGGCTGGAAGGATGCTGCACTGT 2100
Qy 241 CCCGAGGCCCAAGTATGAGGTGAGGCTGCAAGGCTTTCGCAAAAGGACATGCGCGGCACTGTG 2160
Db 2101 CCCGAGGCCCAAGTATGAGGTGAGGCTGCAAGGCTTTCGCAAAAGGACATGCGCGGCACTGTG 2220
Qy 301 ATGTGGCCCGGAGCAGTATGAGGCTTCTAAGTGTGCTGAGTGGGGCGGCGGCTGCTCCCT 360
Db 2161 ATGTGGCCCGGAGCAGTATGAGGCTTCTAAGTGTGCTGAGTGGGGCGGCGGCTGCTCCCT 2220
Qy 361 ACCAGCTCTCCGTGCAAGGCTTCCGCGCTTTCGCAACAGGACATGCGCGGCGGCTGTG 420
Db 2221 ACCAGCTCTCCGTGCAAGGCTTCCGCGCTTTCGCAACAGGACATGCGCGGCGGCTGTG 2280
Qy 421 TGGATGCGGCTGGGGCGGCGGCTTTCGAGTGTGAGCTTCCGAGACCCACCTCC 480
Db 2281 TGGATGCGGCTGGGGCGGCGGCTTTCGAGTGTGAGCTTCCGAGACCCACCTCC 2340
Qy 481 ACCTGCACTGGTAAAGGATGAGTGGATGAGCTGGGCCACTCCGCTGAGCGCTTCTTTCAGG 540
Db 2341 ACCTGCACTGGTAAAGGATGAGTGGATGAGCTGGGCCACTCCGCTGAGCGCTTCTTTCAGG 2400
Qy 541 AGGATGTGGGAGCGGCGCACCGGCTGGTGGCAGCACAGTCCACAGGAGGATGAGGCA 600
Db 2401 AGGATGTGGGAGCGGCGCACCGGCTGGTGGCAGCACAGTCCACAGGAGGATGAGGCA 2460
Qy 601 CCTACTCTCCGCGTGGGGCGGAGTCTGTGGACTTCCGGCTCCGGCTCTCTGAGGCCA 660
Db 2461 CCTACTCTCCGCGTGGGGCGGAGTCTGTGAGCTTCCGGCTCCGGCTCTCTGAGGCCA 2520
Qy 661 AGGTGGTGTCTTAAGGAGCAGTGGCAGCAGGAGGCTGCAAGGAGGAGGAGGAGGCA 720
Db 2521 AGGTGGTGTCTTAAGGAGCAGTGGCAGCAGGAGGCTGCAAGGAGGAGGAGGAGGCA 2580
Qy 721 GTGCCACACTGAGCTGGGAGTGGCCCGGCGGAGGCTGCAAGGAGGAGGAGGAGGATG 780
Db 2581 GTGCCACACTGAGCTGGGAGTGGCCCGGCGGAGGCTGCAAGGAGGAGGAGGAGGATG 2640
Qy 781 GGAAGAAGCTGAGCTCCAGCTCGAAGTGTGAGTGGAGGCCACAGGCTGCAAGGAGGAGG 840
Db 2641 GGAAGAAGCTGAGCTCCAGCTCGAAGTGTGAGTGGAGGCCACAGGCTGCAAGGAGGAGG 2700
Qy 841 TGGTGTGACAGGAGGCGGAGGATGCGGGAGGATGCTGCGAGGCTGGGGG 898
Db 2701 TGGTGTGACAGGAGGCGGAGGATGCGGGAGGATGCTGCGAGGCTGGGGG 2758

RESULT 5
US-10-027-632-272054
; Sequence 272054, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129

	;	LENGTH: 578	
	;	TYPE: DNA	
	;	ORGANISM: Human	
	US-10-027-632-272054		
	Query Match	19.1%; Score 191; DB 17; Length 578;	
	Best Local Similarity	89.2%; Pred. No. 1.8e-39;	
	Matches 206; Conservative	0; Mismatches 25; Indels 0; Gaps 0	
Qy	650	CTCTGAGCCCAAGGTGGTGTGTTGCTAAGGAGCAGCTGGCAGCAGGAAGCTGCAGGCAGA	709
Dd	347	CCCAGAGCCCAAGGTGGTGTGTTGCCAAGGAGCAGCCACACAGGGAGGTTCAGGCTGA	406
Qy	710	GGCAGAGCCAGATGCGCACACTGAGCTCTCGAGGTGGCCCCAGGCCCAGACGAGGTGACCTG	769
Dd	407	GGCGGGGGCCAGTGCCACCGCTGAGCTCGAGGTGGCCCCAGGCCCAGACAGAGGTGACCTG	466
Qy	770	GTAACAAGATGGGAAGAAGCTGAGCTCCAGCTCGAAAAGTGTGCATGGAGGCCACACAGGCTG	829
Dd	467	GTACAAGATGGGAAGAAGCTGAGTTCAGCTCGAAAAGTGGCGGTGGAGGCCGTGGGCTG	526
Qy	830	CACGCGCAGGCTGGTGTGTGCAGCAGGCAGGCCAGGCGGATGCCGGGGAGTA	880
Dd	527	CACACGGAGCTGGTGTGTGCAGCAGGCGGGCCAGGCAGAGGCCGGGGAGTA	577
 RESULT 7			
	US-09-726-643-22		
	; Sequence 22, Application US/09726643		
	; Patent No. US2002028449A1		
	; GENERAL INFORMATION:		
	; APPLICANT: Ruben et al.		
	; TITLE OF INVENTION: 26 Human secreted proteins		
	; FILE REFERENCE: PZ040P1		
	; CURRENT APPLICATION NUMBER: US/09/726,643		
	; CURRENT FILING DATE: 2000-12-01		
	; PRIOR APPLICATION NUMBER: PCT/US00/15187		
	; PRIOR FILING DATE: 2000-06-02		
	; PRIOR APPLICATION NUMBER: 60/137,725		
	; PRIOR FILING DATE: 1999-06-07		
	; NUMBER OF SEQ ID NOS: 190		
	; SOFTWARE: PatentIn Ver. 2.0		
	; SEQ ID NO 22		
	; LENGTH: 1645		
	; TYPE: DNA		
	; ORGANISM: Homo sapiens		
	US-09-726-643-22		
	Query Match	8.8%; Score 88.4; DB 9; Length 1645;	
	Best Local Similarity	46.5%; Pred. No. 4.3e-13;	
	Matches 359; Conservative	0; Mismatches 406; Indels 6; Gaps 2	
Qy	165	GAGGAGCGCAGCTGTGGCTGAGCTGTCCAGATCAGGCTGGGCTGTGACCTGGCTGAAG	224
Dd	581	GAGCGGGTGGTGTGCTGACTTGTGAGCTCTCAAGGGTGGACTTCCCGGCACCTGGTACAG	640
Qy	225	GATGTCGCGACACTGTCCCGAGGCCCAAAGTATGAGGTGCAGGCATCGGCCGGCGCGG	284
Dd	641	GATGGCGAAGGTGGAGGAGAGCGAGTTGCTGGTGGTGAAGATGGATGGCGCGCAACAC	700
Qy	285	GTGCTCCTCTGGGAGATGTGGCCGGGAGCATGACAGGCCTCTACGATGGGTTCAGCGCG	344
Dd	701	CGTCTGATCCTGCTGAGGCCAAAGTCCAGGACAGTGGCGAGTTTGTAGTCAGGACAGAA	760
Qy	345	GGGGGGCGCATTCGCTTACCAGTCTTCGCTGCAAGGCCTCGCGCGCTTTCTGCACAAGGAC	404
Dd	761	GGGGTCTTCGGCCCTTCTTCGGCGGTCACTGTCCAGATCCTCCCGTGCACATCGTGACCCC	820
Qy	405	ATGGCGGGCAGCTGTGTGGATGCGGTGGCTGGGGCCCGCGCGAGTTTGTAGTGTGAGACC	464
Dd	821	CGAGAACATGTGTTCGTGTCATGCCATAACTTCCGAGTGTGTCACTGCTGGCCTGTGAGGTG	880
Qy	465	TCCGAAGCCCAAGTCCACGTGCATGGGTACAAGGATGGCATGGAGCTGGGCGACCTCCCGT	524

Db 881 GACCGAGAGGACGCCCTCTGTGGTTGGTGTACAAAGACGGGCGAGGAGTGTGAGGAGA---GT 937
Qy 525 GAGCGCTTCTTGGCAGGAGATGTGGGACGGGACCGGCTGTGGCAGCCACAGTACAC 584
Db 938 GACTTGTGTGCTGTGAGAAATGAGGGGCCCATCGCCGCTGTGGTCTGCCCGCCACCCAT 997
Qy 585 AGGAGGATGAAGGACACCTACTCTCTGCGCGTGGGCGAGGACTCTGTGGAATTCGGGCTC 644
Db 998 CCCTCAGACGGGGCGAGTTTCAGTGTGCTGTGAGATGAGTGTGCCCTACTTCACTGTC 1057
Qy 645 CGCTCTCTAGCCCCAAGGTGGTTTCTAAGAGACAGTGTGGCAGCGAGAAAGTGTGAG 704
Db 1058 ACCATACAGACGCTCTCTCTGTGGATCG---TGTATCCAGCGGCAAGGTGTATGTGGCA 1114
Qy 705 GCAGAGGAGGAGCCAGTGCACACTGAGCTGCAGGTGGCCAGGCGCCAGCCAGCGAGGTG 764
Db 1115 GCGTGGCCCTGGAGCGTGTGGTGTGACCTGTGAGCTATGCCGCGCTGGGCGAGAGGTG 1174
Qy 765 ACGTGGTACAAAGGATGGGAAGACTGAGCTCCAGCTCGAAAGTGTGATGGAGGCCACA 824
Db 1175 CGCTGGACCAAGGATGGAGAGGTTGTGGAGAGCCCGCGCTCTCTGTGCAAGAGAA 1234
Qy 825 GGCTGCAAGCCAGGCTGTGTGTGAGCAGGACGCGGCGAGTGGCGGAGATATAGC 884
Db 1235 GACACTGTCTCCGCGCTGTGTGCTGCCGCTGTCCAGCTTCAGAGACTCCGGCGAGTACTTG 1294
Qy 885 TGCAGAGCTGGGGCCAGCGGCTCTCTCTTCATCTGGATGTCAAGAGCC 934
Db 1295 TGTGAATTTGACGATGAGTGGCTCTCTTCACTGTACCGTTCACAGAGTC 1344

RESULT 8

US-10-042-141-22
; Sequence 22, Application US/10042141
; Publication No. US20020183503A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/10/042,141
; PRIOR FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/726,643
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-141-22

Query Match 8.8%; Score 88.4; DB 13; Length 1645;
Best Local Similarity 46.5%; Pred. No. 4.3e-13;
Matches 358; Conservative 0; Mismatches 406; Indels 6; Gaps 2;

Qy 165 GAGGAGCGCAGCTGTGGCTGTGAGCTGTCAAGTACAGGCTGGCGTGTGAGCTGTGAAG 224
Db 581 GAGCGGTGTGTGCTGACTTGTGAGCTCTCAAGGGTGGACTTCCCGGCAACCTGGTACAAG 640
Qy 225 GATGGTGCACACTGTCTCCAGGCCCAAGTATAGGTGAGGATCGCGCGGCGG 284
Db 641 GATGGCAGAGGTGGAGGAGGAGCTGTGCTGTGTGAGATGAGTGGCGCAAAAC 700
Qy 285 GTGCTCTCTGTGCAGATGTGGCCCGGACGATGAGGCTCTACGAGTGGCTCAGCGC 344
Db 701 CGTCTGATCTCTGCTGAGGCCAAGTCCAGACAGTGGCGAGTTTGTGATGACAGCAGAA 760
Qy 345 GGGGGCCGCATCGCCCTACAGCTCTCCGTGCAAGGCCCTCGCGCGCTTTCTTGCACAAGGAC 404

Db 761 GGGGTCTCGGCTTCTTTCGGCGTCACTGTCCAAGATCCTCCCGTGCACATCGTGGACCCC 820
Qy 405 ATGCGCGGCGCAGCTGTGTGGATGCCGTGGGCGGCCCGGCGAGTTTGTAGTGTGAGACC 464
Db 821 CGAGAAATGTGTTCGTGCAATGAGCCCTTAACCTCCAGTGTGTCACTGTGGCTGTGAGGTG 880
Qy 465 TCCGAAGCCCAACGCTCCACGTCACCTGGTACAAAGATGGCATGGAGCTGGGCCACTCCCGT 524
Db 881 GACCGAGAGGACGCCCTCTGTGGTTGGTACAAAGACGGGCGAGGAGGTGGAGGAGA---GT 937
Qy 525 GAGCGCTTCTTTCAGAGAGGATGTGGGACCGGGCACCGGCTGGTGGAGCCACAGTACACC 584
Db 938 GACTTCTGTGTGTGTGAGAAATGAGGGGCCCATCGCGCGCTGTGTGCTGCCCGCCACCCAT 997
Qy 585 AGGAGGATGAAGCACCTACTCTCTCGCGGTGGCGGAGGACTCTGTGGAATTCGGGCTC 644
Db 998 CCCTCAGACGGGGCGAGTTTTCAGTGTGCTGTGAGATGAGTGTGCTCTTCACTGTC 1057
Qy 645 CGCGTCTCTGAGCCCAAGGTGGTGTGTTTGAAGAGCAGCTGGCAGCGAGAAAGTGTGAG 704
Db 1058 ACCATACAGACGCTCTCTCTGTGGATCG---TGTATCCAGCGGCAAGGTGTATGTGCA 1114
Qy 705 GCAGAGCAGAGGACCACTGACACTGAGCTGTGAGGTGGCCAGGCCACAGCGAGAGGTG 764
Db 1115 GCGGTGCGCTTGGAGCGTGTGTGTGCTGACCTGTGAGCTATGCGCGCGCTTGGGCGAGAGGTG 1174
Qy 765 ACGTGGTACAAAGGATGGGAAGAGCTGAGCTCCAGCTCGAAAGTGTGATGGAGGCCACA 824
Db 1175 CGCTGGACCAAGGATGGAGAGGTTGTGGAGAGGCCCGCGCTCTCTGTGCAAGAGAA 1234
Qy 825 GGCTGCACGCGCGCAGGCTGTGTGTCAGCAGCGGCGGATGCCGGGAGTATAGC 884
Db 1235 GACACTGTCTCCGCGCTGTGTGCTGCCGCTGTCCAGCTTCAGAGACTCCGGCGAGTACTTG 1294
Qy 885 TGCAGAGCTGGGGCCAGCGGCTCTCTTCCATCTGGATGTCAAGAGCC 934
Db 1295 TGTGAATTTGACGATGAGTGGCGCTCTTCACTGTCAACCGTTCACAGAGTC 1344

RESULT 9

US-10-919-272-22
; Sequence 22, Application US/10919272
; Publication No. US20050010042A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/10/919,272
; CURRENT FILING DATE: 2004-08-17
; PRIOR APPLICATION NUMBER: US/09/726,643
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-919-272-22

Query Match 8.8%; Score 88.4; DB 19; Length 1645;
Best Local Similarity 46.5%; Pred. No. 4.3e-13;
Matches 358; Conservative 0; Mismatches 406; Indels 6; Gaps 2;

Qy 165 GAGGAGCGCAGCTGTGGCTGTGAGCTGTGAGTCAAGTGGCTGGCTGTGAGCTGTGAAG 224
Db 581 GAGCGGTGTGTGCTGACTTGTGAGCTCTCAAGGGTGGACTTCCCGGCAACCTGGTACAAG 640
Qy 225 GATGGTGCACACTGTCTCCAGGCCCAAGTATAGGTGAGGATCGCGCGGCGGCGG 284

Db 641 GATGGGAGAGGTGGAGGAGCGAGTTGCTGTGGTGTGAGATGGATGGCGCAACAC 700
Qy 285 GTGCTCCTTGTGCAGATGTGGCCCGGACCAATGAGCCCTCTACGATGGGTGAGCGGC 344
Db 701 CGTCTGATCTCTGCTGAGGCGCAAGAGTCAGGACAGTGGCGAGTTTGAAGTCAGGACAGAA 760
Qy 345 GGGGGCGCATCGCTTACCAGCTCTCCGTGCAAGCCCTCGCGCGCTTCTGCACAAGGAC 404
Db 761 GGGGCTCTCGGCGCTTCTTCCGCGTCACTGTCCAAGATCTCCCGTGCAATCGTGGACCCC 820
Qy 405 ATGCGCGGCGAGCTGTGTGGATGCGCGTGGGCGCGCGCGAGCTTTGAGTGTGAGAGTG 880
Db 821 CGAGAACATGTGTTCGTGCATGCCATAACTTCCGAGTGTGTCACTGCTGGCCCTGTGAGGTG 880
Qy 465 TCCGAAGCCACGTCCACGTGCACTGTGTACAAGATGGCATGTGAGCTGGGCCCATCTCCGGT 524
Db 881 GACCGAGAGGACGCCCTCTGCTGTGGTGTGTACAAAGACGGGCGAGGAGGTGGAGGAGA--GT 937
Qy 525 GAGCGCTTCTGCAAGGAGGATGTGGGACGGGACCGGCTGTGGTGGGACGACAGTCACC 584
Db 938 GACTTCTGTGTGTGGAGATGAGGGGCGCCCATCGCGCGCTGGTGTGCGCGCCACCCAT 997
Qy 585 AGGAGGATGAAGGACACTACTCTCGCGCGTGGCGGAGGACTCTGTGAGACTTCCGGCTC 644
Db 998 CCCTCAGACGGGGCGAGTTTCAGTGGTGTGCTGGAGATGATGTGCTACTTCACTGTC 1057
Qy 645 CGGCTCTCTGAGCCCAAGGTGGTGTGCTTAAGAGAGAGCTGGACGCGAGAACTGCGAG 704
Db 1058 ACCATCAGACAGCTCTCTCTGTGGATCG--TGTATCCAGCGGCAAGGTGTATGTGCA 1114
Qy 705 GCAGAGCAGAGCCAGTGCACACTGAGTGCAGGTGGCGGAGTGGCCAGGCCAGAGGTG 764
Db 1115 GCGGTGCGCTGGAGCGTGTGGTGTGCTGACCTGTGAGCTATGCCGCGCCCTGGGCGAGAGGTG 1174
Qy 765 ACGTGGTACAGGATGGAGAAAGCTGAGCTCCAGCTCGAAAGTGTGCATGGAGGCCACA 824
Db 1175 CGCTGGACCAAGATGAGAGGAGGTGGTGGAGAGCCCGCGCTGCTCTGCAGAGGAA 1234
Qy 825 GGCTGCACGCGCAGCGTGTGTGTCAGCAGCGCAGGCGGATGCCGGGAGTATAGC 884
Db 1235 GACACTGTCCGCGCGCTGGTGTGCTCCGCTGTCAGCTCGAGGACTCCGGCGAGTACTTG 1294
Qy 885 TCGAGAGCTGGGGCCAGCGCTCTCTTCCATCTGAGTGTCAAGAGCC 934
Db 1295 TGTGAATTGACGATGAGTGGCGCTCTCTTCACTGTCAACCGTCAAGAGTC 1344

RESULT 10

US-10-264-049-850
; Sequence 850, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:

; APPLICANT: Birex et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 850
; LENGTH: 2170
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-264-049-850

Query Match 8.8%; Score 88.4; DB 17; Length 2170;
Best Local Similarity 46.5%; Pred. No. 4,3e-13;
Matches 358; Conservative 0; Mismatches 406; Indels 6; Gaps 2;

Qy 165 GAGGAGGCGCAGCTGCTGGCTGAGCTGTCAAGTCAAGGATCCGCGCGCGTGTGACCTGTGTAAG 224
Db 1072 GAGCGGTGTGTGCTGACTTGTGAGCTCTCAAGGTTGACTTCCCGGCAACCTTGTGTACAG 1131
Qy 225 GATGGTCCGACACACTGTCTCCCGAGGCCCAAGTATGAGGTGCAAGGATCCGCGCGCGCGCGG 284
Db 1132 GATGGGCAAGGTGTGGAGGAGCGAGTTGCTGTGTGTGAAGATGGATGGGCGCAACAC 1191
Qy 285 GTGCTCCTTGTGCAGATGTGGCCCGGACGATCGAGGCTCTACGAGTGGTGTGAGCGCGC 344
Db 1192 CGTCTGATCTCTGCTGAGGCCAAAGTCCAGGACAGTGGCGAGTTTGAAGTGCAGGACAGAA 1251
Qy 345 GGGGGCGCATCGCTTACCAGCTCTCGGTGCAAGCCCTCGCGCGCTTCTTCTGCACAAGGAC 404
Db 1252 GGGGCTCTCGGCTTCTTCCGCGTCACTGTCCAAGATCTCTCCCGTGCAATCGTGGACCCC 1311
Qy 405 ATGCGCGGCGAGCTGTGTGGATGCGCTGGGCGCGCGCGGCGAGTTTGAAGTGTGAGAGC 464
Db 1312 CGAGAACATGTGTTCGTGTCATGCCATAACTTCCGAGTGTGTATGCTGGCCTGTGAGGTG 1371
Qy 465 TCCGAAGCCACGTCCACGTGCACTGTGTACAAGATGGCATGTGAGCTGGGCCACTCCGGT 524
Db 1372 GACCGAGAGGACGCCCTCTGCTGTGGTGTACAAAGACGGGCGAGGAGTGGAGGAGA--GT 1428
Qy 525 GAGCGCTTCTGCAAGGAGGATGTGGGACGGGACCGGCTGTGGCTGTGGCAGCCACAGTCACC 584
Db 1429 GACTTCTGTGTGTGGAGATGAGGGGCGCCCATCGCGCGCTGGTGTGCTCCCGCACCCAG 1488
Qy 585 AGGAGGATGAAGGACACTACTCTCGCGCGTGGCGGAGGAGTCTGTGAGACTTCCGGCTC 644
Db 1489 CCCTCAGACGGGGCGAGTTTCAGTGGTGTGCTGGAGATGATGTGCTACTTCACTGTC 1548
Qy 645 CGGCTCTCTGAGCCCAAGGTGGTGTGCTTAAGAGAGAGCTGGCAGCGAGAGTGTGAG 704
Db 1549 ACCATCAGACAGCTCTCTCTGTGGATCG--TGTATCCAGCGGCAAGGTGTATGGCA 1605
Qy 705 GCAGAGCAGGAGCCAGTGCACACTGAGTGCAGGTGGCGGCGGCGGCGGCGGCGGAGGTG 764
Db 1606 GCGGTGCGCTGGAGCGTGTGGTGTGCTGACCTGTGAGCTATGCCGCGCTGGGCGAGAGGTG 1665
Qy 765 ACGTGGTACAGGATGGAGAAAGCTGAGCTCCAGCTCGAAAGTGTGATGGAGGCCACA 824
Db 1666 CGCTGGACCAAGATGAGAGGAGGTGGTGGAGAGCCCCGCGCTGCTCTGCAGAGGAA 1725
Qy 825 GGCTGCACGCGCAGGCTGGTGTGTCAGCAGCAGGCGGCGGATGCCGGGAGTATAGC 884
Db 1726 GACACTGTCCGCGCGCTGGTGTGCTCCGCTGTCCAGCTCGAGGACTCCGCGGAGTACTTG 1785

RESULT 11

US-10-108-260A-1946
; Sequence 1946, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:

; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: HI-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1946
; LENGTH: 2534
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-108-260A-1946

Query Match 8.6%; Score 86.2; DB 17; Length 2534;

Db 493 GCTTTGAGTGGCTCTCTGTCCACAGAGAGTGCCAGCGACCCGGGCCATGTGGACAGTCCGGTG 552
Qy 781 GGAAGAGAGTGTAGCTCCAGCTCGAAGAGTGTGCATGGAGGCCACAGGCTGCACGGCGAGGC 840
Db 553 GGAAGACAGTGGGAGCTTCAGCGCGCTTCAGGCGCCACAGTCAAGGCGCGAATAATACATCC 612
Qy 841 TGGTTGTGACAGCGCAGCGCAGCGGATGCGCGGGAGTATAGTGCAGAGGCTGGGGGCC 900
Db 613 TGGTGGTCCGGAGGCTGCACCAAGTGTATGCCGGGAGGTGGTCTTCTGTGCGGGGCC 672
Qy 901 AGCGGCTCTCTTCCATCTGTGATGTCAAGAGCCCAAGGTGGTGTGTTGCCAAGACAGG 960
Db 673 TCACCTCAAGGCTCACTCACTTGTCAAGAGAGCGCGCGCCATCATCAAGCCCTGG 732
Qy 961 TGGCACACAGTGTAGTGCAGGCTCAGGCGAGGG 993
Db 733 AGACCAGTGGTGGCGCCAGGAGGACGTGG 765

RESULT 14

US-10-093-463-71

; Sequence 71, Application US/10093463

; Publication No. US20030208039A1

; GENERAL INFORMATION:

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Shenoy, Suresh

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Gusev, Vladimir

; APPLICANT: Pochart, Pascal

; APPLICANT: Zhong, Mei

; APPLICANT: Rastelli, Luca

; APPLICANT: Mezes, Peter

; APPLICANT: Smithson, Glennda

; APPLICANT: Guo, Xiaojia

; APPLICANT: Gerlach, Valerie

; APPLICANT: Casman, Stacie

; APPLICANT: Boldog, Ferenc

; APPLICANT: Li, Li

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Tchernev, Velizar

; APPLICANT: Gangolli, Esha

; APPLICANT: Vernet, Corine

; APPLICANT: Pena, Carol

; APPLICANT: Burgess, Catherine

; APPLICANT: Liu, Xiaohong

; APPLICANT: Spytek, Kimberly

; APPLICANT: Gorman, Linda

; APPLICANT: Spaderna, Steven

; APPLICANT: Voss, Edward

; APPLICANT: Malyankar, Uriel

; APPLICANT: Anderson, David

; APPLICANT: Patturajan, Meera

; APPLICANT: Miller, Charles

; APPLICANT: Taupier, Raymond J. Jr.

; TITLE OF INVENTION: No. US20030208039A1el Antibodies that Bind to Antigenic Polypepti

; TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.

; FILE REFERENCE: 21402-290A (Cura 590AT)

; CURRENT APPLICATION NUMBER: US/10/093,463

; PRIOR FILING DATE: 2002-06-24

; PRIOR FILING DATE: 2001-04-14

; PRIOR FILING DATE: 2001-04-14

; PRIOR FILING DATE: 2001-12-03

; PRIOR FILING DATE: 2001-03-08

; PRIOR FILING DATE: 2001-03-08

; PRIOR FILING DATE: 2001-09-27

; PRIOR FILING DATE: 2001-07-10

; PRIOR APPLICATION NUMBER: 60/279,995

; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 71
; LENGTH: 14109
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15) .. (14088)
US-10-093-463-71

Query Match 7.9% Score 78.6; DB 17; Length 14109;

Best Local Similarity 48.5%; Pred. No. 1.4e-10;

Matches 278; Conservative 0; Mismatches 289; Indels 6; Gaps 2;

Qy 421 TGGATGCCGTGGGCGCCGCGCAGTTTGAAGTGTGAGACCTCCGAAGCCACGTC 480
Db 199 TGTCCGACAGGAGCGCGCACCTGGCCCTGCAGTGTGAAGTCTCTGACCCGAGGCC 258
Qy 481 ACCTGACCTGTACAGGATGGCATGGAGCTGGGCCACTCGGTGAGCGCTTCTTGAGG 540
Db 259 ATGTGTGTGGCGCAAGATGGCGTGCAGCTGGGCCCCAGTGACATGACTTCTGTC 318
Qy 541 AGGATGTGGGACGCGCACCGGCTGTGTGCAGCCACAGTCACCAGGACGATGAAGGCA 600
Db 319 ACAGGCGGACACGCGG---GGCTGTGTGATGACGTGAGCCCTGAGACGCCGCGCC 375
Qy 601 CCTACTCTCCGCGTGGGCGAGGACTCTGTGGAATTCCGGCTCCGCTCTCTGAGGCCA 660
Db 376 TGTACACCTGCCAGTGGGCTCCGAGGAGACCCGGGCGCGGTCCGCTGCAGATCTGC 435
Qy 661 AGGTGTGTTTGTAAAGGACAGCTGCGACGCGAGGAAGCTGCGAGGACGAGGACGCA 720
Db 436 ACCTGGGCATCACCAGAGGCTGAAAGACA---ATGGAGGTGCTGGAAGGGGAAAGCTGCA 492
Qy 721 GTGCCACACTGAGCTGGAGGTGSCCCAGGCCAGACGGAGGTGACGTGGTGTACAAGATG 780
Db 493 GCTTTGAGTGGCTCTGTCTCCACAGAGAGTCCAGCGACCCGCGCCATGTGGACAGT 552
Qy 781 GGAAGAGCTGAGCTCCAGCTCGAAGTGTGCATGGAGGCCACAGGTGACGCGCGCAGGC 840
Db 553 GGAAGACAGTGGGCGAGCTCCAGCGCTTCCAGGCCACACGTCAGGCGCGAATAATACATCC 612
Qy 841 TGGTTGTGACAGGCGAGGCCAGCGGATCGGGGAGTATAGCTGCGAGGCTGGGGGCC 900
Db 613 TGGTGGTCCGGAGGCTGCAAGTGTATCGGGGAGGTGGTCTTCTGTGCGGGGCC 672
Qy 901 AGCGGCTCTCTTCCATCTGATGTCAAGAGCCCAAGGTGGTGTGTTGCCAAGACGAGG 960
Db 673 TCACCTCAAGGCTCACTCATTTGTACAGAGAGGCGCGCCCATCATCAAGCCCTGG 732

Qy 961 TGGCACAGTGAGTGCAGCTGAGCGAGGG 993
Db 733 AAGACCACTGGTGGCGCCAGGGAGACGTGG 765

RESULT 15

US-09-726-643-44
; Sequence 44, Application US/09726643
; Patent No. US2002028449A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/09/726,643
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 2254
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-643-44

Query Match 6.8%; Score 67.8; DB 9; Length 2254;
Best Local Similarity 49.7%; Pred. No. 8.5e-08;
Matches 256; Conservative 0; Mismatches 252; Indels 7; Gaps 3;

Qy 420 GTGATGCCGTGGCTGGGGCCCGCGCAGCTTTGAGTGTGAGACCTCCGAAGCCACGTC 479
Db 1331 GTGATGCCAACTTCCGAGTGTGTCATGCTGGCCTGTGAGGTGACCCGAGAGACGCC 1390
Qy 480 CACGTGCACTGTCACAGGATGGCATGAGCTGGGCCACTCCGGTGAGGGCTTCTTGCGAG 539
Db 1391 CCTGTGCTTGGTACAGGACGGGAGGAGTGG---AGGAGTGTGACTTCTGTTGCTG 1447
Qy 540 GAGGATGTGGGAGCGCGCACCGGTGGTGGCAGCCACAGTCACCAGGCAGGATGAAGGC 599
Db 1448 GAGAATGAGGGGCCCCATCGCGCTGTGTGCCCGCCACCAGCCCTCAGACGGGGGC 1507
Qy 600 ACCTACTCCTCGCGGTGGGCGAGGACTCTGTGAGACTTCGGGCTCCGGCTCTCTGAGCCC 659
Db 1508 GAGTTTCAGTCGCTGCGTGGAGATGAGTGTGCTACTTCACTGTCACTCATCACAGACGTC 1567
Qy 660 AAGTGTGTGTTGCTAAGGAGCAGCTGGCAGCAGAGCTGCAGGCAGAGCAGGAGCC 719
Db 1568 TCCTCGTGGATCGT---GTATCCAGCGGCAAGGTGTATGTGGCAGCCGTGGCCTGGAG 1624
Qy 720 AGTGCACACTGAGCTGCGAGGTGGCCCGAGGCCAGAGGAGTACGTTGGTACAAAGGAT 779
Db 1625 CGTGTGGTGTGACCTGTGAGCTATGCCGGCCCTGGGCGAGAGTGGCTGGACCAAGGAT 1684
Qy 780 GGGAGAGAGCTGAGCTCCAGCTCGAAAGTGTGCAATGGAGGCCACAGGCTGCACGCGCAGG 839
Db 1685 GGAGAGGAGGTGGTGGAGACCCCG-CGCTGCTCTCTGCAGAGGAAGACACTGTTCGCGCGC 1743
Qy 840 CTGCTTGTGACGAGGAGCGGAGCGGATGCGGGGAGTATAGCTGCGAGGCTGGGGGC 899
Db 1744 CTGCTGCTGCGCGGTGTTCAGCTCGAGGACTCCGGCGAGTACTTGTGTGAAATTGACGAT 1803
Qy 900 CAGCGGCTCTCTTCCATCTGGATGTCAAAGAGCC 934
Db 1804 GAGTCGGCTCTTCACTGTCACTGTCACTGTCACTGTCACTGTCACTGTCACTGTCACTGT 1838

Search completed: March 21, 2005, 16:07:49
Job time : 631.052 secs

THIS PAGE BLANK (USPTO)

	Db	 510	CCACCGCTGGTCA CGGGGCTGTGAGCAGTTTGAGACC CGCACAGA CCTCATCTCTCAT	569
	Qy	 121	CCTGGAGCTGTGCTCATCCGAGGAGCTGTGTGAACCGCTGTATACAGAAAGGGCGTGGTCAC	180
	Db	 570	CCTGGAGCTGTGCTCATCCGAGGAGCTGTGTGAACCGCTGTATACAGAAAGGGCGTGGTGAC	629
	Qy	 181	GGAGGCCGAGGTCAAAGTCTA CATCAGCAGCTGTGTGGAGGGGCTGCATCACTTGCA CAG	240
	Db	 630	GGAGGCCGAGGTCAAAGTCTA CATCAGCAGCTGTGTGGAGGGGCTGCATCACTTGCA CAG	689
	Qy	 241	CCATGGCGTTCTCCACCTGGACAATAAGCCCTTAACATCCTGATGTGTGATCCTTGCCCG	300
	Db	 690	CCATGGCGTTCTCCACCTGGACAATAAGCCCTTAACATCCTGATGTGTGATCCTTGCCCG	749
	Qy	 301	GGAAGACATTAAAATCTGCGACTTTGGCTTTTGCCCGAGAACATCACCCAGCAGAGCTGCA	360
	Db	 750	GGAAGACATTAAAATCTGCGACTTTGGCTTTTGCCCGAGAACATCACCCAGCAGAGCTGCA	809
	Qy	 361	GTTCAGCAGTACGGCTCCCCTGAGTTCGTCTCCCGCAGATCATCCAGCAGAAACCCCTGT	420
	Db	 810	GTTCAGCAGTACGGCTCCCCTGAGTTCGTCTCCCGCAGATCATCCAGCAGAAACCCCTGT	869
	Qy	 421	GAGCGAAGCCTCCGACATTTGGGCCATGGTGTATCTCTACCTCAGCCTGACCTGCTC	480
	Db	 870	GAGCGAAGCCTCCGACATTTGGGCCATGGTGTATCTCTACCTCAGCCTGACCTGCTC	929
	Qy	 481	ATCCCCATTTGCGGCGAGAGTAGCGTGCACCTCCTGAAACGTCCTGGAGGGGCGCGT	540
	Db	 930	ATCCCCATTTGCGGCGAGAGTAGCGTGCACCTCCTGAAACGTCCTGGAGGGGCGCGT	989
	Qy	 541	GTCTAGGACAGCCCACTGCTGCCCACTCAGCGAAGAGCCCAAAGACTTTCATCAAGGC	600
	Db	 990	GTCTAGGACAGCCCACTGCTGCCCACTCAGCGAAGAGCCCAAAGACTTTCATCAAGGC	1049
	Qy	 601	TACGCTCAGAGAGCCCTCAGGCGCGGCTAGTGGCGGCCAGTGCCTCTCCCAACCCCTG	660
	Db	 1050	TACGCTCAGAGAGCCCTCAGGCGCGGCTAGTGGCGGCCAGTGCCTCTCCCAACCCCTG	1109
	Qy	 661	GTTCCTGAAATCCATGCTCGCGAGGAGGCCCATCTCATCAACCAAGCAGAGCTCAAGTT	720
	Db	 1110	GTTCCTGAAATCCATGCTCGCGAGGAGGCCCATCTCATCAACCAAGCAGAGCTCAAGTT	1169
	Qy	 721	CCTCTGGCCCGAAGTGCCTGGCAGCGTTCCTGATGAGCTACAAGTCCAATCCTGGTGAT	780
	Db	 1170	CCTCTGGCCCGAAGTGCCTGGCAGCGTTCCTGATGAGCTACAAGTCCAATCCTGGTGAT	1229
	Qy	 781	GCAGCTCCATCCCTCAGCTGCTGGGGGCCCAACCGCAGAGCCCTCCCTCGCGGTAGCCCG	840
	Db	 1230	GCAGCTCCATCCCTCAGCTGCTGGGGGCCCAACCGCAGAGCCCTCCCTCGCGGTAGCCCG	1289
	Qy	 841	GCACCTCTGAGGAGACACTGGTGGCTCCTCCAGTTCCTCTCTCTCTCTGACAAACAGCT	900
	Db	 1290	GCACCTCTGAGGAGACACTGGTGGCTCCTCCAGTTCCTCTCTCTCTGACAAACAGCT	1349
	Qy	 901	CGCCCCATTTGCCCGGCTAAGTCACTGCCACCTCCCGGTGACACACTCAACCACTGCT	960
	Db	 1350	CGCCCCATTTGCCCGGCTAAGTCACTGCCACCTCCCGGTGACACACTCAACCACTGCT	1409
	Qy	 961	GCACCCCGGGGCTTCTGCGGCCCTCGGCAGGCTGCTGAGGAAAGCCGAGGCCAGTGA	1020
	Db	 1410	GCACCCCGGGGCTTCTGCGGCCCTCGGCAGGCTGCTGAGGAAAGCCGAGGCCAGTGA	1469
	Qy	 1021	GCAGCTCCACGAGGCCCAAGCTCCGCTGATCTCTCCGAGGGTCCGGGCCACCGGCCCG	1080
	Db	 1470	GCAGCTCCACGAGGCCCAAGCTCCGCTGATCTCTCCGAGGGTCCGGGCCACCGGCCCG	1529
	Qy	 1081	CCAGGGGTGCTGCCCGGCA GCGTCA TCCCGAGCCTGTTTACCAACAGGCGGGTGA	1140
	Db	 1530	CCAGGGGTGCTGCCCGGCA GCGTCA TCCCGAGCCTGTTTACCAACAGGCGGGTGA	1589
	Qy	 1141	GAGCCTTGAGCAGGGGCTTGCCCGGGAGCAGGCGGCACCCCGCCGGCGCGGCA	1200

Db	1590	GAGCCTGAGACA	CGGGGGCCCTGGCC	CCCGGGAGAGAGCGGCA	CCCGGGCCGGCGCGCA	1649
QY	1201	CTTGCTGAAGGG	CGGGTACATTTGCGGGGGCGCT	GCCAGGCGCTTGC	GCAGCACTGATGGA	1260
Db	1650	CCTGCTGAAGGG	CGGGTACATTTGCGGGGGCGCT	TGCCAGGCTTGC	GCAGGCCACTGATGA	1709
QY	1261	GCACCGGCTGCT	TGGAGGAGAGGCCCGC	AGGAGGAGAGCGCC	TCTTGCGCCAAAGC	1320
Db	1710	GCACCGGCTGCT	TGGAGGAGGAGCGCGC	AGGAGGAGAGCGCC	TCTTGCGCCAAAGC	1769
QY	1321	CCCTCATTTGAG	ACTGGCCCTCGGCTGCT	TGCTTGGACCCACT	TGCGCCCTGGGCA	1380
Db	1770	CCCTCATTTGAG	ACTGGCTGCTTGGCTTGGAC	CCACTTGGCCCTTGGCCCA	1829	
QY	1381	CAGCCATCTCC	TGGAAATGACTCTCGAG	CACCCCGCCCTCTCTCGAG	GGCCTGCGG	1440
Db	1830	CAGCCATCTCC	TGGAAATGACTCTCGAG	CACCCCGCCCTCTCTCGAG	GGCCTGCGG	1889
QY	1441	TGAGGCACAG	CAGCTGCTTACAGCCCCT	TCGGGGGGGCGCC	TATCAGGGA	1500
Db	1890	TGAGGCACAG	CACTGCTTACAGCCCCT	TCGGGGGGGCGCC	TATCAGGGA	1949
QY	1501	CCCTCAGGCT	CCAAAGCATTCTCTCGAG	CACCCCGCCCTCTCTCGAG	GGCCTGCGG	1560
Db	1950	CCCTCAGGCT	CCAAAGCATTCTCTCGAG	CACCCCGCCCTCTCTCGAG	GGCCTGCGG	2009
QY	1561	GAGGCGATCCC	CGGACAGCCCTTGGGGG	CAGCCAGCCCTTCTGCG	ACCCCAAGAGG	1620
Db	2010	GAGGCGATCCC	CGGACAGCCCTTGGGGG	CAGCCAGCCCTTCTGCG	ACCCCAAGAGG	2069
QY	1621	TTCTGCCCCC	CAGGAGGCTG	CAGCCCCACCCAGCAGT	TGCCCATGCGCTCTTGGCTC	1680
Db	2070	TTCTGCCCCC	CAGGAGGCTG	CAGCCCCACCCAGCAGT	TGCCCATGCGCTCTTGGCTC	2129
QY	1681	CTTCCCTCAG	ATCTTGC	AAAGGCCCTTAGTA	CCCTCAAGCCCTTCTTGGGACA	1740
Db	2130	CTTCCCTCAG	ATCTTGC	AAAGGCCCTTAGTA	CCCTCAAGCCCTTCTTGGGACA	2189
QY	1741	GCCCCAGG	CACCCCTG	CCCAAAGCCCCCAT	TGGACTCTAAGATGGGGCC	1800
Db	2190	GCCCCAGG	CACCCCTG	CCCAAAGCCCCCAT	TGGACTCTAAGATGGGGCC	2249
QY	1801	TGGAGACATCT	CTTCTG	TGGAGGCCAAACCCGGGCCCT	TGCAAGTCCCAGGGTCAGC	1860
Db	2250	TGGAGACATCT	CTTCTG	TGGAGGCCAAACCCGGGCCCT	TGCAAGTCCCAGGGTCAGC	2309
QY	1861	CTCCAGGCG	AGCTCTTCC	CAAGTGAGCTCCCTCAG	GGTGGCTCTCCAGGTGGGCAC	1920
Db	2310	CTCCAGGCG	AGCTCTTCC	CAAGTGAGCTCCCTCAG	GGTGGCTCTCCAGGTGGGCAC	2369
QY	1921	AGAGCTGGC	CCCCCTCTG	GATGCGAGGGTGG	ACCCAGGAGGCTGAGGATCTGTCCGA	1980
Db	2370	AGAGCTGGC	CCCCCTCTG	GATGCGAGGGTGG	ACCCAGGAGGCTGAGGATCTGTCCGA	2429
QY	1981	CTCCACACCA	CCCTTG	CAGCGCCCTCAGGA	ACAGGACCAATGCGAAGTCTCCCTGGG	2040
Db	2430	CTCCACACCA	CCCTTG	CAGCGCCCTCAGGA	ACAGGACCAATGCGAAGTCTCCCTGGG	2489
QY	2041	TGGTCGGGG	GGGTACG	CAGCGCTGGCTG	GCTATGGCACTTTGCCCTTTGGTGGAGATGC	2100
Db	2490	TGGTCGGGG	GGGTACG	CAGCGCTGGCTG	GCTATGGCACTTTGCCCTTTGGTGGAGATGC	2549
QY	2101	AGGGGGCA	TGTTGGGG	CAGGGCCATGTTGGGC	CAGATAGCTTGGGCTGTGTCCCAGTC	2160
Db	2550	AGGGGGCA	TGTTGGGG	CAGGGCCATGTTGGGC	CAGATAGCTTGGGCTGTGTCCCAGTC	2609
QY	2161	GGAGGAGAG	GAGCAGG	AGGCGCAGGCTT	CCAGTCCAGTTCGAGGAGCAGAGGAGGC	2220
Db	2610	GGAGGAGAG	GAGCAGG	AGGCGCAGGCTT	CCAGTCCAGTTCGAGGAGCAGAGGAGGC	2669
QY	2221	CAGGGCTG	AGAGCCCA	CTGCCAGGTCAGT	GCAAGGCCCTGTGCCCTGAGGTCGCGAGGGC	2280
Db	2670	CAGGGCTG	AGAGCCCA	CTGCCAGGTCAGT	GCAAGGCCCTGTGCCCTGAGGTCGCGAGGGC	2729


```
QY 2281 TCCACACGAGGCTCTCCAGAGCCCAACCCCATGGAGGACATCGGGAGGCTCTCCCTGGT 2340
Db |||||
QY 2730 TCCACACGAGGCTCTCCAGAGCCCAACCCCATGGAGGACATCGGGAGGCTCTCCCTGGT 2789
Db |||||
QY 2341 GCAGATCCGGACCTGTCAAGTATGCGAGGGCGGACACAAATATCCCTGGACATTTTC 2400
Db |||||
QY 2790 GCAGATCCGGACCTGTCAAGTATGCGAGGGCGGACACAAATATCCCTGGACATTTTC 2849
Db |||||
QY 2401 CGAGGTGGACCCCGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
Db |||||
QY 2850 CGAGGTGGACCCCGCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2909
Db |||||
QY 2461 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGGCAGAGCGCCCTCCCCAT 2520
Db |||||
QY 2910 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGGCAGAGCGCCCTCCCCAT 2969
Db |||||
QY 2521 GGCTGAGGAGAGCTGCCGAGTTCCCGGAGCCCAAGTCCGCTCAGGCAGAGCGCCCTCCCCAT 2580
Db |||||
QY 2970 GGCTGAGGAGAGCTGCCGAGTTCCCGGAGCCCAAGTCCGCTCAGGCAGAGCGCCCTCCCCAT 3029
Db |||||
QY 2581 CCCCCACGAGGCTGGAGATCAAGAGGAGTCAAGAGGATGTGACCGCTGTGGCAGA 2640
Db |||||
QY 3030 CCCCCACGAGGCTGGAGATCAAGAGGAGTCAAGAGGATGTGACCGCTGTGGCAGA 3089
Db |||||
QY 2641 GGCTGCCGTGGCAGGAGCGCAAGTGGTCTCGCCCTCAGCGAGCTCTTCCACTTCCC 2700
Db |||||
QY 3090 GGCTGCCGTGGCAGGAGCGCAAGTGGTCTCGCCCTCAGCGAGCTCTTCCACTTCCC 3149
Db |||||
QY 2701 TGGGAGGACCTGCCGCTGGATGAGCTGCAGAGCTGGGGCTGGTGAGAGGTGAAGGC 2760
Db |||||
QY 3150 TGGGAGGACCTGCCGCTGGATGAGCTGCAGAGCTGGGGCTGGTGAGAGGTGAAGGC 3209
Db |||||
QY 2761 CTCGCTGGAGCACATCTCCCGGATCTTGAAGGGCAGGCCGGAAGTCTGGAGAGGAGGG 2820
Db |||||
QY 3210 CTCGCTGGAGCACATCTCCCGGATCTTGAAGGGCAGGCCGGAAGTCTGGAGAGGAGGG 3269
Db |||||
QY 2821 GCGCCCGCAGAGAGGAGCGGCTTGTCTTCTCCGCTCTCAGGTCTGAAGAGCTGGGA 2880
Db |||||
QY 3270 GCGCCCGCAGAGAGGAGCGGCTTGTCTTCTCCGCTCTCAGGTCTGAAGAGCTGGGA 3329
Db |||||
QY 2881 CCAGAGCCGACATCTCTAAGGAGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCACT 2940
Db |||||
QY 3330 CCAGAGCCGACATCTCTAAGGAGCTCTCAGATGAGACTGTGTCTCTGGGCCAGTCACT 3389
Db |||||
QY 2941 GACACTGGCTGCGCAGGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGGAGCAAGACGG 3000
Db |||||
QY 3390 GACACTGGCTGCGCAGGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGGAGCAAGACGG 3449
Db |||||
```

RESULT 2

```
US-10-077-130-3
; Sequence 3, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; FILE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047P1RCP1(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7893
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-10-077-130-3
```

```
Query Match 100.0%; Score 3001; DB 13; Length 7893;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 CTTACGAGCAGAACTCTGGGCCCCAGGCATACAGGAGCGAGACATCTCTGGCCGCGCTGAG 60
Db |||||
QY 3486 CTTACGAGCAGAACTCTGGGCCCCAGGCATACAGGAGCGAGACATCTCTGGCCGCGCTGAG 3545
Db |||||
QY 61 CCACCCGCTGCTCAACCGGGCTGTGGACCACTTTGAGACCCGCAAGACCTCTCATCTCAT 120
Db |||||
QY 3546 CCACCCGCTGCTCAACCGGGCTGTGGACCACTTTGAGACCCGCAAGACCTCTCATCTCAT 3605
Db |||||
QY 121 CTTGAGCTGTGCTCATCTCGAGGAGCTGTGGAACCGCTGTACAGGAAGGCGGTGGTGAAC 180
Db |||||
QY 3606 CTTGAGCTGTGCTCATCTCGAGGAGCTGTGGAACCGCTGTACAGGAAGGCGGTGGTGAAC 3665
Db |||||
QY 181 GGAGGCCGAGGTCAGGCTCTACATCCAGCAGCTGGTGGAGGGCTGCACCTACCTGCACAG 240
Db |||||
QY 3666 GGAGGCCGAGGTCAGGCTCTACATCCAGCAGCTGGTGGAGGGCTGCACCTACCTGCACAG 3725
Db |||||
QY 241 CCATGGGCTTCTCACCTGGACATAAAGCCCTCTAAACATCTCTGATGGTGCATCTCTGCCG 300
Db |||||
QY 3726 CCATGGGCTTCTCACCTGGACATAAAGCCCTCTAAACATCTCTGATGGTGCATCTCTGCCG 3785
Db |||||
QY 301 GGAAGACATTTAAATCTGCGACTTTTGGCTTTGCCGGAACATCAACCCAGCAGAGCTGCA 360
Db |||||
QY 3786 GGAAGACATTTAAATCTGCGACTTTTGGCTTTGCCGGAACATCAACCCAGCAGAGCTGCA 3845
Db |||||
QY 361 GTTCAGCAGTACGGCTCCCTCGAGTTCTCTCCCGAGATCATCCAGCAGAAACCTGT 420
Db |||||
QY 3846 GTTCAGCAGTACGGCTCCCTCGAGTTCTCTCCCGAGATCATCCAGCAGAAACCTGT 3905
Db |||||
QY 421 GAGCAGAGCTCTCGACATTTGGGCTATGGGCTGTCTCTCTACCTCAGCTGACCTGCTC 480
Db |||||
QY 3906 GAGCAGAGCTCTCGACATTTGGGCTATGGGCTGTCTCTCTACCTCAGCTGACCTGCTC 3965
Db |||||
QY 481 ATCCCCATTTGCCGGCAGAGTACCGTGCCACCTCTCTGAAACGTCTCTGAGGGGCGCT 540
Db |||||
QY 3966 ATCCCCATTTGCCGGCAGAGTACCGTGCCACCTCTCTGAAACGTCTCTGAGGGGCGCT 4025
Db |||||
QY 541 GTCATGAGCAGAGCCCATCTGCTGCGCCACCTCAGGGAAGACGCCAAGACTTCATCAAGGC 600
Db |||||
QY 4026 GTCATGAGCAGAGCCCATCTGCTGCGCCACCTCAGGGAAGACGCCAAGACTTCATCAAGGC 4085
Db |||||
QY 601 TAGCTGACAGAGCCCTCAGAGCCGCGCTAGTGGCGCCAGTGGCTCTCCCAACCCCTG 660
Db |||||
QY 4086 TAGCTGACAGAGCCCTCAGAGCCGCGCTAGTGGCGCCAGTGGCTCTCCCAACCCCTG 4145
Db |||||
QY 661 GTTCTCTGAAATCCATGCTGCGGAGGAGGCCCTTTCATCAACCAAGCAGCTCAAGTT 720
Db |||||
QY 4146 GTTCTCTGAAATCCATGCTGCGGAGGAGGCCCTTTCATCAACCAAGCAGCTCAAGTT 4205
Db |||||
QY 721 CTTCTGCGCCGAGAGTCTGCTGCGAGCGTTCCCTGATGAGCTCAAGTTCATCTCTGGTGTAT 780
Db |||||
QY 4206 CTTCTGCGCCGAGAGTCTGCTGCGAGCGTTCCCTGATGAGCTCAAGTTCATCTCTGGTGTAT 4265
Db |||||
QY 781 GCGCTCCATCCCTGAGCTGTGCGGGGCCACCGCAGAGCCCTCTCCCTCGGCGTAGGCCG 840
Db |||||
QY 4266 GCGCTCCATCCCTGAGCTGTGCGGGGCCACCGCAGAGCCCTCTCCCTCGGCGTAGGCCG 4325
Db |||||
QY 841 GCACCTCTGAGGAGACACTGCTGCTCCTCAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 900
Db |||||
QY 4326 GCACCTCTGAGGAGACACTGCTGCTCCTCAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4385
Db |||||
QY 901 CGCCCATTTTCCCGGCTTAAGTCACTGCCACCTCTCCCGGTGACACTCACCACCTGCT 960
Db |||||
QY 4386 CGCCCATTTTCCCGGCTTAAGTCACTGCCACCTCTCCCGGTGACACTCACCACCTGCT 4445
Db |||||
QY 961 GCACCCCGGGGCTTCTCTGCGGCCCTCTGCGCAGCTGCGCTGAGGAAGCCGAGGCGAGTGA 1020
Db |||||
```


Db 4446 GCACCCCGGGCTTCCTGCGGCCCTCGGCCAGCTGCTGAGGAAGCCGAGGCCAGTGA 4505
Qy 1021 GCGCTCCACGAGGCCAGCTCCGCTCGCATCTCCGAGGCTGCGGGCCACCGGCCG 1080
Db 4506 GCGCTCCACGAGGCCAGCTCCGCTCGCATCTCCGAGGCTGCGGGCCACCGGCCG 4565
Qy 1081 CCAGGGCTGCGTCCCGGCCACAGGCTCATCCGCAAGCTGTCTACCAACAGGCGGGTGA 1140
Db 4566 CCAGGGCTGCGTCCCGGCCACAGGCTCATCCGCAAGCTGTCTACCAACAGGCGGGTGA 4625
Qy 1141 GAGCCCTGAGACAGGGGCCCTTGCGCCCGGGGAGAGGCGGACCCGGCCCGGGCGGCA 1200
Db 4626 GAGCCCTGAGACAGGGGCCCTTGCGCCCGGGGAGAGGCGGACCCGGCCCGGGCGGCA 4685
Qy 1201 CCGTGTGAAGGCGGGTACATTGCGGGGCGCTGCGAGGCCCTGCGGAGCCACTGATGGA 1260
Db 4686 CCGTGTGAAGGCGGGTACATTGCGGGGCGCTGCGAGGCCCTGCGGAGCCACTGATGGA 4745
Qy 1261 GCACCGCTGTGAGAGGAGGCGCGCAGGAGGAGAGGCCACCTCTCTGGCCAAAGC 1320
Db 4746 GCACCGCTGTGAGAGGAGGCGCGCAGGAGGAGAGGCCACCTCTCTGGCCAAAGC 4805
Qy 1321 CCCCTCATTTCCAGACTGCGCTCCGCTGCGCTGCTGCGACCCCGCCCTCTCGAGGCCCTGCGG 1380
Db 4806 CCCCTCATTTCCAGACTGCGCTCCGCTGCGCTGCTGCGACCCCGCCCTCTCGAGGCCCTGCGG 4865
Qy 1381 CAGCACTCTCTGAAACATGACTCTCCGAGACACCCCGCCCTCTCTCGAGGCCCTGCGG 1440
Db 4866 CAGCACTCTCTGAAACATGACTCTCCGAGACACCCCGCCCTCTCTCGAGGCCCTGCGG 4925
Qy 1441 TGAGGCAACAGGACTGCTTTCAGCGCCCTCGGGGGGCGCCCTATCAGAGGACATGAGGCA 1500
Db 4926 TGAGGCAACAGGACTGCTTTCAGCGCCCTCGGGGGGCGCCCTATCAGAGGACATGAGGCA 4985
Qy 1501 CCCTCAGGGCTCCAAAGAGCTTCCATCACTGTGTGGCCACCCAGGCACTGCTCAGCCAGA 1560
Db 4986 CCCTCAGGGCTCCAAAGAGCTTCCATCACTGTGTGGCCACCCAGGCACTGCTCAGCCAGA 5045
Qy 1561 GAGGCCATTCGCCGACAGCCCTTGGGGGCGACCCAGCCCTTTCGCAACCCCAAGCAGGG 1620
Db 5046 GAGGCCATTCGCCGACAGCCCTTGGGGGCGACCCAGCCCTTTCGCAACCCCAAGCAGGG 5105
Qy 1621 TTCTGCCCCCAGAGAGGCTGACGCCCCAGCCAGAGTTCGCCCATGCCCTCTGCTC 1680
Db 5106 TTCTGCCCCCAGAGAGGCTGACGCCCCAGCCAGAGTTCGCCCATGCCCTCTGCTC 5165
Qy 1681 CTTCCCTCCAGGATCTTTGCAAGAGGCCCCCTTAGTACCCTCAAGCCCCCTTCTTGGGACA 1740
Db 5166 CTTCCCTCCAGGATCTTTGCAAGAGGCCCCCTTAGTACCCTCAAGCCCCCTTCTTGGGACA 5225
Qy 1741 GCCCCAGGACCCCTGCCCCCTGCGCAAGCAAGCCCCCCTTGGACTCTAAGATGGGGC 1800
Db 5226 GCCCCAGGACCCCTGCCCCCTGCGCAAGCAAGCCCCCCTTGGACTCTAAGATGGGGC 5285
Qy 1801 TGGAGACATCTCTTCTGAGAGGCCAAACCCGCGCCCTGAGTTCCCGAGGTCAGC 1860
Db 5286 TGGAGACATCTCTTCTGAGAGGCCAAACCCGCGCCCTGAGTTCCCGAGGTCAGC 5345
Qy 1861 CTCCAGGAGGATCTTCTCCAGTGAAGCTCCCTCAGGGTGGGCTCCTCCAGGTGGGGC 1920
Db 5346 CTCCAGGAGGATCTTCTCCAGTGAAGCTCCCTCAGGGTGGGCTCCTCCAGGTGGGGC 5405
Qy 1921 AGAGCTGGCCCTCCCTGAGTGGGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 1980
Db 5406 AGAGCTGGCCCTCCCTGAGTGGGAGGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 5465
Qy 1981 CTCACACCCACCTTCAGCGGGCTCAGGAACAGGCGACCATGCGCAAGTTCTCCCTGGG 2040
Db 5466 CTCACACCCACCTTCAGCGGGCTCAGGAACAGGCGACCATGCGCAAGTTCTCCCTGGG 5525
Qy 2041 TGTCTGGGGGGTACGCAAGGCTGCTGGCTATGGCACTTTGCTTTGGTGGAGATGC 2100
Db 5526 TGTCTGGGGGGTACGCAAGGCTGCTGGCTATGGCACTTTGCTTTGGTGGAGATGC 5585

Qy 2101 AGGGGGCATCTCTGGGCGAGGGGCCCATGTGGGCCAGGATAGCTTGGGCTGTGTCCAGTC 2160
Db 5586 AGGGGGCATCTCTGGGCGAGGGGCCCATGTGGGCCAGGATAGCTTGGGCTGTGTCCAGTC 5645
Qy 2161 GAGGAGGAGGAGCAGGAGGAGGCCAGGGCTGAGTCCAGTCTCGAGAGGAGCAGCAGGAGGC 2220
Db 5646 GAGGAGGAGGAGCAGGAGGAGGCCAGGGCTGAGTCCAGTCTCGAGAGGAGCAGCAGGAGGC 5705
Qy 2221 CAGGGCTGAGAGGCCACTGCCCCAGGTCAGTGCAAGGCTGTGTCTGTAGGTCTGGCAGGGC 2280
Db 5706 CAGGGCTGAGAGGCCACTGCCCCAGGTCAGTGCAAGGCTGTGTCTGTAGGTCTGGCAGGGC 5765
Qy 2281 TCCACACAGGAGCTCTCCAGAGCCACCCCATGGGAGGACATCGGCAAGTCTTCCCTGGT 2340
Db 5766 TCCACACAGGAGCTCTCCAGAGCCACCCCATGGGAGGACATCGGCAAGTCTTCCCTGGT 5825
Qy 2341 GCAGATCCGGGACCTGTCAAGTGTATCGGAGGCGCGCACAAATATCCCTGGACATTTTC 2400
Db 5826 GCAGATCCGGGACCTGTCAAGTGTATCGGAGGCGCGCACAAATATCCCTGGACATTTTC 5885
Qy 2401 CGAGGTGGACCCCGCTACTCAACCTCTCAGACCTGTAGATATCAAGTACCTCCCAT 2460
Db 5886 CGAGGTGGACCCCGCTACTCAACCTCTCAGACCTGTAGATATCAAGTACCTCCCAT 5945
Qy 2461 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCCGTCAAGCCAGAGCCGCTTCCCAT 2520
Db 5946 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCCGTCAAGCCAGAGCCGCTTCCCAT 6005
Qy 2521 GCGTGAAGGAGGCTGCGCGAGTTCCCGAGCCCAAGTGGCCCTGCGCAAGTGAATGGG 2580
Db 6006 GCGTGAAGGAGGCTGCGCGAGTTCCCGAGCCCAAGTGGCCCTGCGCAAGTGAATGGG 6065
Qy 2581 CCCCCAGCAGGCTGGAGATCACAGAGAGTCAAGAGTGTGGACGCGCTGCTGGCAGA 2640
Db 6066 CCCCCAGCAGGCTGGAGATCACAGAGAGTCAAGAGTGTGGACGCGCTGCTGGCAGA 6125
Qy 2641 GCGTCCGTGGGCGAGGAAGCGCAAGTGTCTCGCCGTCAAGCAGCTCTTCCACTTCCC 2700
Db 6126 GCGTCCGTGGGCGAGGAAGCGCAAGTGTCTCGCCGTCAAGCAGCTCTTCCACTTCCC 6185
Qy 2701 TGGGAGGACCTGCGCGCTGGATGAGCTGCAAGCTGGGGCTGCTGAGAGTGAAGGC 2760
Db 6186 TGGGAGGACCTGCGCGCTGGATGAGCTGCAAGCTGGGGCTGCTGAGAGTGAAGGC 6245
Qy 2761 CTCCTGAGAGCACAATCTCCCGATCTGAAAGGCGAGGCGGAAGGTCTGGAAGAGGAGG 2820
Db 6246 CTCCTGAGAGCACAATCTCCCGATCTGAAAGGCGAGGCGGAAGGTCTGGAAGAGGAGG 6305
Qy 2821 GCCCCCAGGAAGAAAGCCAGGCTTGTCTTCCCTTCCGGCTCTCAGGTCTGAAGAGCTGGGA 2880
Db 6306 GCCCCCAGGAAGAAAGCCAGGCTTGTCTTCCCTTCCGGCTCTCAGGTCTGAAGAGCTGGGA 6365
Qy 2881 CCGAGGCGGACATTCCTAAGGAGCTCTCAGATGAGACTGTGGTCTGGGCGACTCAGT 2940
Db 6366 CCGAGGCGGACATTCCTAAGGAGCTCTCAGATGAGACTGTGGTCTGGGCGACTCAGT 6425
Qy 2941 GACACTGCGCTGCCAGGTGTCAAGCCAGCAGCTGCCAGGCGACCTCGAGCAAGAGCGG 3000
Db 6426 GACACTGCGCTGCCAGGTGTCAAGCCAGCAGCTGCCAGGCGACCTCGAGCAAGAGCGG 6485
Qy 3001 A 3001
Db 6486 A 6486

RESULT 3
US-10-077-130-1
; Sequence 1, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.

; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family

; FILE REFERENCE: MIP2001-047P1RCP1(M)

; CURRENT APPLICATION NUMBER: US/10/077,130

; PRIOR FILING DATE: 2002-02-15

; PRIOR FILING DATE: 2001-02-15

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 8106

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: 5'UTR

; LOCATION: (1)...(71)

; NAME/KEY: CDS

; LOCATION: (72)...(7964)

; NAME/KEY: 3'UTR

; LOCATION: (7965)...(8106)

; US-10-077-130-1

Query Match 100.0%; Score 3001; DB 13; Length 8106;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	CCTACGGAGCAGAACTCGGGCCCAAGGATACAGGGAGCGAGACATCTCTGGCCGCGCTGAG	60
Db	3557	CCTACGGAGCAGAACTCGGGCCCAAGGATACAGGGAGCGAGACATCTCTGGCCGCGCTGAG	3616
Qy	61	CCACCCGCTGTCACGGGGCTGTCGACCAAGTTTGAGACCGCAAGACCCCTCATCTTCAT	120
Db	3617	CCACCCGCTGTCACGGGGCTGTCGACCAAGTTTGAGACCGCAAGACCCCTCATCTTCAT	3676
Qy	121	CCTGGAGCTGTCTCATCCAGGAGCTGCTGGACCGCTGTACAGGAAGGCGTGGTGAC	180
Db	3677	CCTGGAGCTGTCTCATCCAGGAGCTGCTGGACCGCTGTACAGGAAGGCGTGGTGAC	3736
Qy	181	GGAGCCGAGTGTAAGGTCTACATCCAGAGCTGGTGAGGGGTGCACTACCTGCGACAG	240
Db	3737	GGAGCCGAGTGTAAGGTCTACATCCAGAGCTGGTGAGGGGTGCACTACCTGCGACAG	3796
Qy	241	CCATGGGCTTCTCACCTGGACATAAAGCCCTCTAAACATCTGTATGTGTGATCTGCGCG	300
Db	3797	CCATGGGCTTCTCACCTGGACATAAAGCCCTCTAAACATCTGTATGTGTGATCTGCGCG	3856
Qy	301	GGAGACATTAATAATCTGCGACTTTGGCTTTGGCCAGAAATCAACCCAGCAGAGCTGCA	360
Db	3857	GGAGACATTAATAATCTGCGACTTTGGCTTTGGCCAGAAATCAACCCAGCAGAGCTGCA	3916
Qy	361	GTTCAGCCAGTACCGGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGCAGAAACCTGT	420
Db	3917	GTTCAGCCAGTACCGGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGCAGAAACCTGT	3976
Qy	421	GAGCGAGCCTCCGACATTTGGGCGCATGGGTGATCTCTACCTACGCTGACCTGCTC	480
Db	3977	GAGCGAGCCTCCGACATTTGGGCGCATGGGTGATCTCTACCTACGCTGACCTGCTC	4036
Qy	481	ATCCCCCATTTGCGCGCAGAGTGACCGTGCCACCTCTCTGAACCTCTCTGAGGGGGCGGT	540
Db	4037	ATCCCCCATTTGCGCGCAGAGTGACCGTGCCACCTCTCTGAACCTCTCTGAGGGGGCGGT	4096
Qy	541	GTATGAGCAGCCCCATGGCTGCCCACTTCAGCGAAGACGCCAAGACTTTCATCAAGGC	600
Db	4097	GTATGAGCAGCCCCATGGCTGCCCACTTCAGCGAAGACGCCAAGACTTTCATCAAGGC	4156
Qy	601	TACGCTGAGAGAGCCCTCAGGCCCGGCTAGTGGCGGCCAGTGCTCTCCACCCCTG	660
Db	4157	TACGCTGAGAGAGCCCTCAGGCCCGGCTAGTGGCGGCCAGTGCTCTCCACCCCTG	4216
Qy	661	GTTCCTGAATTCATGCTCGGAGAGGGCCCACTTCATCAACACCAGCAGCTCAAGTT	720
Db	4217	GTTCCTGAATTCATGCTCGGAGAGGGCCCACTTCATCAACACCAGCAGCTCAAGTT	4276

Qy	721	CTCTCTGGCCCGAAGTCGCTGGCAGCGGTTCCCTGATGAGCTACAAGTCCATCTCTGGTGAT	780
Db	4277	CTCTCTGGCCCGAAGTCGCTGGCAGCGGTTCCCTGATGAGCTACAAGTCCATCTCTGGTGAT	4336
Qy	781	GGCTCTCATCCCTGAGCTGTGGGGGCCCAACCCGACAGCCCTCCCTCGGCGTAGCCCG	840
Db	4337	GGCTCTCATCCCTGAGCTGTGGGGGCCCAACCCGACAGCCCTCCCTCGGCGTAGCCCG	4396
Qy	841	GCACCTCTGACGGAGACACTGGTGGCTCCTCAAGTTCTCTCTCTCTCTCTCTGACACAGACT	900
Db	4397	GCACCTCTGACGGAGACACTGGTGGCTCCTCAAGTTCTCTCTCTCTCTCTCTGACACAGACT	4456
Qy	901	CGCCCCATTTTCCCGGGCTAAGTCACTGCCACCTCCCGGTGACACACTCACCACTGCT	960
Db	4457	CGCCCCATTTTCCCGGGCTAAGTCACTGCCACCTCCCGGTGACACACTCACCACTGCT	4516
Qy	961	GCACCCCGGGGGCTTCTCTGGGGCTTCTCTGGGGCTTCTCTGGGGCTTCTCTGGGGCTTCT	1020
Db	4517	GCACCCCGGGGGCTTCTCTGGGGCTTCTCTGGGGCTTCTCTGGGGCTTCTCTGGGGCTTCT	4576
Qy	1021	GGCTTCCACCGAGCCCGAGCTCCGCTGATCTCCCGAGGGTCCCGGGCCACCCGGCCGC	1080
Db	4577	GGCTTCCACCGAGCCCGAGCTCCGCTGATCTCCCGAGGGTCCCGGGCCACCCGGCCGC	4636
Qy	1081	CGAGGCTGCTGCTCCCGGGCAAGCTCCGAGCTCATCCGAGCTTCTTACCAACAGGGGGTGA	1140
Db	4637	CGAGGCTGCTGCTCCCGGGCAAGCTCCGAGCTCATCCGAGCTTCTTACCAACAGGGGGTGA	4696
Qy	1141	GAGCCCTGAGCAGCGGGCCCTGGCCCGGGGAGCAGCGCGCACCCGGCCCGGGCCCGGCA	1200
Db	4697	GAGCCCTGAGCAGCGGGCCCTGGCCCGGGGAGCAGCGCGCACCCGGCCCGGGCCCGGCA	4756
Qy	1201	CTGTGTAAGGGCGGTACATTTGGGGGGCGCTGCGAGGCTGCGAGGACCTATGATGGA	1260
Db	4757	CTGTGTAAGGGCGGTACATTTGGGGGGCGCTGCGAGGCTGCGAGGACCTATGATGGA	4816
Qy	1261	GCACCGGCTCTGAGAGAGAGCGCCGACGGGAGGAGCAGCCACCTCTCTGGGCCAAGC	1320
Db	4817	GCACCGGCTCTGAGAGAGAGCGCCGACGGGAGGAGCAGCCACCTCTCTGGGCCAAGC	4876
Qy	1321	CCCTCATTTGAGACTGCTCCGCTGCTGCTGCTGCGACCCACCTTGGCCCCCTGGCCA	1380
Db	4877	CCCTCATTTGAGACTGCTCCGCTGCTGCTGCTGCGACCCACCTTGGCCCCCTGGCCA	4936
Qy	1381	CAGCCACTCCCTGAGAACATGACTCTCGAGCAGCCCCCGGCCCTCTCTCGAGGCTCGGG	1440
Db	4937	CAGCCACTCCCTGAGAACATGACTCTCGAGCAGCCCCCGGCCCTCTCTCGAGGCTCGGG	4996
Qy	1441	TGAGGCAACAGGACTGCTTACGCCCTCTCGGGGGGGGGCCCTTATCAGGGACATGGGGCA	1500
Db	4997	TGAGGCAACAGGACTGCTTACGCCCTCTCGGGGGGGGGCCCTTATCAGGGACATGGGGCA	5056
Qy	1501	CCCTCAGGGCTCAAGCAGCTTCCATCCTGCTGGGCCACCCAGGCACTGCTCAGGCCAGA	1560
Db	5057	CCCTCAGGGCTCAAGCAGCTTCCATCCTGCTGGGCCACCCAGGCACTGCTCAGGCCAGA	5116
Qy	1561	GAGGCCATCCCGGACAGCCCTTTGGGGGAGCAGCAGCCCTTTTGTGCCACCCCAAGCAGGG	1620
Db	5117	GAGGCCATCCCGGACAGCCCTTTGGGGGAGCAGCAGCCCTTTTGTGCCACCCCAAGCAGGG	5176
Qy	1621	TTCTGCCCCCGAGAGAGGGCTGAGCCGCCACCCAGCAGTTTGGCCCCATGCTCTTGGCTC	1680
Db	5177	TTCTGCCCCCGAGAGAGGGCTGAGCCGCCACCCAGCAGTTTGGCCCCATGCTCTTGGCTC	5236
Qy	1681	CTTCCCTTCAGGATCTTTGCAAGAGGGCCCTTAGTACCTCTCAAGCCCTCTTCTTGGGACA	1740
Db	5237	CTTCCCTTCAGGATCTTTGCAAGAGGGCCCTTAGTACCTCTCAAGCCCTCTTCTTGGGACA	5296
Qy	1741	GCCCCAGGCAACCCCTGCCCCCTGCCAAAGAACGAGCCGCCCTTGGACTCTAAGATGGGGCC	1800
Db	5297	GCCCCAGGCAACCCCTGCCCCCTGCCAAAGAACGAGCCGCCCTTGGACTCTAAGATGGGGCC	5356

QY 1801 TGGAGACATCTCTCTCTGGAGGGGCAAAACCCGGGCCCTTGCAGTTTCCCAAGGTCAGC 1860
Db 5357 TGGAGACATCTCTCTCTGGAGGGGCAAAACCCGGGCCCTTGCAGTTTCCCAAGGTCAGC 5416
QY 1861 CTCCACAGGGAGCTCTTCCCAAGTCAGCTCCCTCAGGGTGGGCTCCTCCAGGTGGGCAC 1920
Db 5417 CTCCACAGGGAGCTCTTCCCAAGTCAGCTCCCTCAGGGTGGGCTCCTCCAGGTGGGCAC 5476
QY 1921 AGAGCTGGCCCTCTCTGGATGCGGAGGGCTGGAACCCAGGAGGCTGAGGATCTGTCCGA 1980
Db 5477 AGAGCTGGCCCTCTCTGGATGCGGAGGGCTGGAACCCAGGAGGCTGAGGATCTGTCCGA 5536
QY 1981 CTCACACCCACCTTGCAGCGGCTTCAGAAACAGGCGAACATGCGCAAGTTCTCCCTGGG 2040
Db 5537 CTCACACCCACCTTGCAGCGGCTTCAGAAACAGGCGAACATGCGCAAGTTCTCCCTGGG 5596
QY 2041 TGGTCGGGGGGCTACGACAGGCGTGGCTTATGGCACCTTTGCTTTGGTGAGATGC 2100
Db 5597 TGGTCGGGGGGCTACGACAGGCGTGGCTTATGGCACCTTTGCTTTGGTGAGATGC 5656
QY 2101 AGGGGGCATGTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGCCAGTC 2160
Db 5657 AGGGGGCATGTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGCCAGTC 5716
QY 2161 GGAGGAGGAGCAGGAGGAGGCGCAGGGCTGAGTCCAGTCGGAGGAGCAGGAGGC 2220
Db 5717 GGAGGAGGAGCAGGAGGAGGCGCAGGGCTGAGTCCAGTCGGAGGAGCAGGAGGC 5776
QY 2221 CAGGCTGAGAGCCACCTGCCCCAGCTCAGTGCAGAGGCTGTGCTCAGGTCGCGAGGC 2280
Db 5777 CAGGCTGAGAGCCACCTGCCCCAGCTCAGTGCAGAGGCTGTGCTCAGGTCGCGAGGC 5836
QY 2281 TCCACACGAGAGCTCTCAGAGGCCACCCCATGGGAGGACATCGGGCAGGTCCTCCCTGGT 2340
Db 5837 TCCACACGAGAGCTCTCAGAGGCCACCCCATGGGAGGACATCGGGCAGGTCCTCCCTGGT 5896
QY 2341 GCAGATCGGAGCTGTGTCAGTGTATGGGAGGCGGCGGACACAAATATCCCTGGACATTC 2400
Db 5897 GCAGATCGGAGCTGTGTCAGTGTATGGGAGGCGGCGGACACAAATATCCCTGGACATTC 5956
QY 2401 CGAGGTGAGCCCGCCCTACCTCAACCTCTCAGACCTGTGACGATATCAAGTACCTCCCAT 2460
Db 5957 CGAGGTGAGCCCGCCCTACCTCAACCTCTCAGACCTGTGACGATATCAAGTACCTCCCAT 6016
QY 2461 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCGCTCAGCCAGAGCGCCCTCCCCAT 2520
Db 6017 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCGCTCAGCCAGAGCGCCCTCCCCAT 6076
QY 2521 GGCTGAGGAGAGCTGGCCGAGTTCCCGAGGCCACCGTGGCCCTGGCCAGGCTGAG 2580
Db 6077 GGCTGAGGAGAGCTGGCCGAGTTCCCGAGGCCACCGTGGCCCTGGCCAGGCTGAG 6136
QY 2581 CCCCACGAGGCTGGAGATCAGAGGAGTCAGAGGATGTGGAACGCGCTGTGGCAGA 2640
Db 6137 CCCCACGAGGCTGGAGATCAGAGGAGTCAGAGGATGTGGAACGCGCTGTGGCAGA 6196
QY 2641 GGCTGCGGTGGCAGGAGCGCAAGTGGTCTCTGCGCTCAGCAGGCTCTTCCACTTCC 2700
Db 6197 GGCTGCGGTGGCAGGAGCGCAAGTGGTCTCTGCGCTCAGCAGGCTCTTCCACTTCC 6256
QY 2701 TGGAGGACCTTGGCGCTGATGAGCTGACAGAGCTGGGCTCGGTGAGAGTGAAGGC 2760
Db 6257 TGGAGGACCTTGGCGCTGATGAGCTGACAGAGCTGGGCTCGGTGAGAGTGAAGGC 6316
QY 2761 CTCGTGGAGCAATCTCCCGATCTTGAAGGCGAGGCCGGAAGGCTTGGAGAGGAGG 2820
Db 6317 CTCGTGGAGCAATCTCCCGATCTTGAAGGCGAGGCCGGAAGGCTTGGAGAGGAGG 6376
QY 2821 GCCCCAGGAAAGCAGGCTTGGTCTTCCGCTCTCAGGCTCTCAGGCTGAAAGCTGGGA 2880
Db 6377 GCCCCAGGAAAGCAGGCTTGGTCTTCCGCTCTCAGGCTCTCAGGCTGAAAGCTGGGA 6436
QY 2881 CCAGGCGCCGACATTCCTAAGGAGGCTCTCAGATGAGCTGTGGTCTCTGGGCGAGTCAGT 2940

Db 6437 CCGAGCGCCGACATTCCTAAGGAGCTCTCAGATGAGACTGTGTCTCTGGCCAGTCAGT 6496
QY 2941 GACACTGGCTGCCAGGTGTGAGCCAGCAGCAGTGGCCAGCCACCTGGAGCAAGACGG 3000
Db 6497 GACACTGGCTGCCAGGTGTGAGCCAGCAGCAGTGGCCAGCCACCTGGAGCAAGACGG 6556
QY 3001 A 3001
Db 6557 A 6557
RESULT 4
US-10-077-130-6
; Sequence 6, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047P1BCP1(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 23907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-077-130-6
Query Match 100.0%; Score 3001; DB 13; Length 23907;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 19500 CCTACGGAGCAGAACTCGGGCCCGAGGCATACAGGGAGCGAGACATCTCGCGCGCTGAG 60
Db 19500 CCTACGGAGCAGAACTCGGGCCCGAGGCATACAGGGAGCGAGACATCTCGCGCGCTGAG 19559
QY 61 CCACCCGCTGGTCCAGGGGCTCTGGACCAAGTTTGGACCCGCAAGACCTCATCTCAT 120
Db 19560 CCACCCGCTGGTCCAGGGGCTCTGGACCAAGTTTGGACCCGCAAGACCTCATCTCAT 19619
QY 121 CTGGAGCTGTCTCATCGAGAGCTGTGACCGCTGTACAGGAAGGGCGTGGTAC 180
Db 19620 CTGGAGCTGTCTCATCGAGAGCTGTGACCGCTGTACAGGAAGGGCGTGGTAC 19679
QY 181 GGAGGCCGAGGTCAGAGTCTACATCCAGCAGCTGGTGGAGGGGCTGCACCTACTGCACAG 240
Db 19680 GGAGGCCGAGGTCAGAGTCTACATCCAGCAGCTGGTGGAGGGGCTGCACCTACTGCACAG 19739
QY 241 CCATGGGCTTCCACCTGGACATAAAGCCCTCTAACAATCTTGATGGTGATCTCTGCCCG 300
Db 19740 CCATGGGCTTCCACCTGGACATAAAGCCCTCTAACAATCTTGATGGTGATCTCTGCCCG 19799
QY 301 GGAAGACATTAAATCTCGACTTTGGCTTTGCCCGAGAACATCACCCAGCAGAGCTGCA 360
Db 19800 GGAAGACATTAAATCTCGACTTTGGCTTTGCCCGAGAACATCACCCAGCAGAGCTGCA 19859
QY 361 GTTCAGCAGTACGCTCCCTCTGAGTGTCTCTCCCGAGATCATCCAGCAAGAACCTGT 420
Db 19860 GTTCAGCAGTACGCTCCCTCTGAGTGTCTCTCCCGAGATCATCCAGCAAGAACCTGT 19919
QY 421 GAGCGAAGCTCCGACATTTGGGCCCATGGGTGTATCTCTTACCTCAGCCTGACCTGTC 480
Db 19920 GAGCGAAGCTCCGACATTTGGGCCCATGGGTGTATCTCTTACCTCAGCCTGACCTGTC 19979
QY 481 ATCCCAATTTGGCGGCGAGAGTGACCGGTGCCACCTCTCTGAACGTCCTGGAGGGGCGGT 540
Db 19980 ATCCCAATTTGGCGGCGAGAGTGACCGGTGCCACCTCTCTGAACGTCCTGGAGGGGCGGT 20039

Qy	541	GTCTGAGGAGCCCCCATGGCTGCCACCTCAGCGAAGACGCGCAAGACTTTCATCAAGGC	600
Db	20404	GTCTGAGGAGCCCCCATGGCTGCCACCTCAGCGAAGACGCGCAAGACTTTCATCAAGGC	20999
Qy	601	TACGCTGAGAGAGCCCTCAGGCGCCGCTAGTGCAGGCGCCAGTGCCTCTCCACCCCTG	660
Db	20100	TACGCTGAGAGAGCCCTCAGGCGCCGCTAGTGCAGGCGCCAGTGCCTCTCCACCCCTG	20159
Qy	661	GTCTCTGAAATCCATGCTGCGGAGGCGCCACTTTCATCAACCAAGCAGCTCAAGTT	720
Db	20160	GTCTCTGAAATCCATGCTGCGGAGGCGCCACTTTCATCAACCAAGCAGCTCAAGTT	20219
Qy	721	CCTCTGCGCCGAGTCCCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTGGTGAT	780
Db	20220	CCTCTGCGCCGAGTCCCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTGGTGAT	20279
Qy	781	CGCTTCCATCCCTGAGCTGCTGCGGGGCCACCCGACAGCCCTCCCTCGGCGTAGCCCG	840
Db	20280	CGCTTCCATCCCTGAGCTGCTGCGGGGCCACCCGACAGCCCTCCCTCGGCGTAGCCCG	20339
Qy	841	GCACCTCTGAGGAGACATGCTGGCTCTCTCAGTTCTCTCTCTCTGACAAGAGCT	900
Db	20340	GCACCTCTGAGGAGACATGCTGGCTCTCTCAGTTCTCTCTCTCTCTGACAAGAGCT	20399
Qy	901	CGCCCATTTGCGCGGCTAAGTCTACTGCGACCTCCCGCGGTGACACACTCACACTGCT	960
Db	20400	CGCCCATTTGCGCGGCTAAGTCTACTGCGACCTCCCGCGGTGACACACTCACACTGCT	20459
Qy	961	GCACCCCGGGCTTCTGCGGCCCTCGGCGAGCTGCTGAGGAAGCGGAGGCGAGTGA	1020
Db	20460	GCACCCCGGGCTTCTGCGGCCCTCGGCGAGCTGCTGAGGAAGCGGAGGCGAGTGA	20519
Qy	1021	CGCTTCCACCGAGGCCAGCTCGCTGCTGATCTCCCGAGGGTGCGGGCCACCGGGCGC	1080
Db	20520	CGCTTCCACCGAGGCCAGCTCGCTGCTGATCTCCCGAGGGTGCGGGCCACCGGGCGC	20579
Qy	1081	CCAGGGCTGCGTCCCGGCGACAGGCTCATCGCGAGCTGTTTACACACGAGGCGGCTGA	1140
Db	20580	CCAGGGCTGCGTCCCGGCGACAGGCTCATCGCGAGCTGTTTACACACGAGGCGGCTGA	20639
Qy	1141	GAGCCCTGAGCACGGGGCCCTGGCCCCGGGAGAGCGGCGCACCGGCCCCGGCGCGCA	1200
Db	20640	GAGCCCTGAGCACGGGGCCCTGGCCCCGGGAGAGCGGCGCACCGGCCCCGGCGCGCA	20699
Qy	1201	CCTGCTGAAGCGGGTACATTTGCGGGGCGCTGCGAGGCTGCGCGAGGCACTGATGGA	1260
Db	20700	CCTGCTGAAGCGGGTACATTTGCGGGGCGCTGCGAGGCTGCGCGAGGCACTGATGGA	20759
Qy	1261	GCACCGCGTGTGAGGAGGAGCGCGCCAGGGAGGAGGAGGCGCACCTCTGGCGCAAGC	1320
Db	20760	GCACCGCGTGTGAGGAGGAGCGCGCCAGGGAGGAGGAGGCGCACCTCTGGCGCAAGC	20819
Qy	1321	CCCTCTATTGAGACTGCGCTCCGGCTGCTGCGACCTGCTGCGACCTTGGGCCCTGGGCA	1380
Db	20820	CCCTCTATTGAGACTGCGCTCCGGCTGCTGCGACCTGCTGCGACCTTGGGCCCTGGGCA	20879
Qy	1381	CAGCCACTCCCTTGGAAACATGACTCTCCGAGCACCCCCCGCCCCCTCTCGGAGGCGTGG	1440
Db	20880	CAGCCACTCCCTTGGAAACATGACTCTCCGAGCACCCCCCGCCCCCTCTCGGAGGCGTGG	20939
Qy	1441	TGAGGCAAGGCACTGCTTCAAGCCCCCTCGGGGGGGCCCCCTATCAGGGAGATGGGGCA	1500
Db	20940	TGAGGCAAGGCACTGCTTCAAGCCCCCTCGGGGGGGCCCCCTATCAGGGAGATGGGGCA	20999
Qy	1501	CCCTCAGGGCTCCAGAGAGTTCATCCACTGGTGGGCCACCCAGGCACTGCTCAGCCAGA	1560
Db	21000	CCCTCAGGGCTCCAGAGAGTTCATCCACTGGTGGGCCACCCAGGCACTGCTCAGCCAGA	21059
Qy	1561	GAGGCCATCCCGGACAGCCCTTGGGGGCGAGCAGCCCTTTCTGCCACCCCAAGCAGGG	1620
Db	21060	GAGGCCATCCCGGACAGCCCTTGGGGGCGAGCAGCCCTTTCTGCCACCCCAAGCAGGG	21119

Qy	1621	TTCTGCCCGCCAGGAGGGCTGACGCCCCCAAGAGTTGCCCATGCCCTCTGGCTC	1680
Db	21120	TTCTGCCCGCCAGGAGGGCTGACGCCCCCAAGAGTTGCCCATGCCCTCTGGCTC	21179
Qy	1681	CTTCCCTCCAGGATCTTGCAGGAGGCGCCCTTAGTACCTCTCAAGCCCTCTCTTGGGACA	1740
Db	21180	CTTCCCTCCAGGATCTTGCAGGAGGCGCCCTTAGTACCTCTCAAGCCCTCTCTTGGGACA	21239
Qy	1741	GCCCCAGGACCCCTGCTGCGCCCAAGAGCCGCCCATTTGAGTCTTAGATGGGGCC	1800
Db	21240	GCCCCAGGACCCCTGCTGCGCCCAAGAGCCGCCCATTTGAGTCTTAGATGGGGCC	21299
Qy	1801	TGGAGACATCTCTCTCTGGAGGSCCAAAACCCGCGCCCTGACAGTTCCTCCAGGGTCAGC	1860
Db	21300	TGGAGACATCTCTCTCTGGAGGSCCAAAACCCGCGCCCTGACAGTTCCTCCAGGGTCAGC	21359
Qy	1861	CTCCACGCGAGCTCTTCCCAAGTGAGCTCCCTCAGGGTGGGCTCTCCAGGTGGGCAC	1920
Db	21360	CTCCACGCGAGCTCTTCCCAAGTGAGCTCCCTCAGGGTGGGCTCTCCAGGTGGGCAC	21419
Qy	1921	AGAGCTTGCGCCCTCTCTGGATGCGGAGGGCTGGACCCAGGAGGTGAGGATCTGTCCGA	1980
Db	21420	AGAGCTTGCGCCCTCTCTGGATGCGGAGGGCTGGACCCAGGAGGTGAGGATCTGTCCGA	21479
Qy	1981	CTCCACACCCACCTTTGAGCGGCTCTCAGGAACAGCGGACCATGCGCAAGTTCTCCCTGG	2040
Db	21480	CTCCACACCCACCTTTGAGCGGCTCTCAGGAACAGCGGACCATGCGCAAGTTCTCCCTGG	21539
Qy	2041	TGGTTCGGGGGGCTTACGAGGCGTGGCTATGGACCTTTGCTTTGGTGGAGATGC	2100
Db	21540	TGGTTCGGGGGGCTTACGAGGCGTGGCTATGGACCTTTGCTTTGGTGGAGATGC	21599
Qy	2101	AGGGGGATGCTGGGGCAGGGGCCCATGTGGGCGAGGATAGCCTGGGCTGTGTCCAGTTC	2160
Db	21600	AGGGGGATGCTGGGGCAGGGGCCCATGTGGGCGAGGATAGCCTGGGCTGTGTCCAGTTC	21659
Qy	2161	GGAGGAGGAGGAGCAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGAGGC	2220
Db	21660	GGAGGAGGAGGAGCAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGAGGC	21719
Qy	2221	CAGGGCTGAGAGCCACTCGCCCGCAGGTGAGTGCAGGCGCTGTGCTGAGGTGCGGAGGC	2280
Db	21720	CAGGGCTGAGAGCCACTCGCCCGCAGGTGAGTGCAGGCGCTGTGCTGAGGTGCGGAGGC	21779
Qy	2281	TCCCAACGAGAGCTCTCCAGAGCCCACTGAGGAGGAGATCGGGCAGGTCTCCCTGGT	2340
Db	21780	TCCCAACGAGAGCTCTCCAGAGCCCACTGAGGAGGAGATCGGGCAGGTCTCCCTGGT	21839
Qy	2341	GCAGATCCGGAGCCTGTCAGGTGATGCGGAGGCGCGACACATATATCCCTGGACATTTTC	2400
Db	21840	GCAGATCCGGAGCCTGTCAGGTGATGCGGAGGCGCGACACATATATCCCTGGACATTTTC	21899
Qy	2401	CGAGGTGAGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTT	2460
Db	21900	CGAGGTGAGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTT	21959
Qy	2461	CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCAGAGCGCGCTCCCCCAT	2520
Db	21960	CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCAGAGCGCGCTCCCCCAT	22019
Qy	2521	GGCTCAGGAGGAGCTGGCCGAGTTCCCGGAGCCCACTGCGGCCCTGGCCAGGTGAACTGGG	2580
Db	22020	GGCTCAGGAGGAGCTGGCCGAGTTCCCGGAGCCCACTGCGGCCCTGGCCAGGTGAACTGGG	22079
Qy	2581	CCCCACGCGGCTGGAGATCACAGAGGAGTTCAGAGGATGTGACGCGCTGTGTCGACAGA	2640
Db	22080	CCCCACGCGGCTGGAGATCACAGAGGAGTTCAGAGGATGTGACGCGCTGTGTCGACAGA	22139
Qy	2641	GGCTGCGGTGGGCAAGGCGCAAGTGTCTTCGCGGTACGCGAGCTCTTTCACCTTCCC	2700
Db	22140	GGCTGCGGTGGGCAAGGCGCAAGTGTCTTCGCGGTACGCGAGCTCTTTCACCTTCCC	22199
Qy	2701	TGGGAGGACCTCGCGCTGGATGAGCCTGACAGACTGGGGCTGGCTGAGAGAGTGAAGGC	2760

Db 22200 TGGGAGGCACTGCGCGCTGGATGAGCTGCGAGAGCTGGGGCTGCGTGAGAGAGTGAAGGC 22259
Qy 2761 CTCCTGTGAGGACATCTCCGGATCTGAAAGGCGAGGCGGGAAGGTCTGGAGAAGGAGG 2820
Db 22260 CTCCTGTGAGGACATCTCCGGATCTGAAAGGCGAGGCGGGAAGGTCTGGAGAAGGAGG 22319
Qy 2821 GCGCCCGAGGAAGAGCCAGGCTGCTTCCTTCCGGCTCTCAGGTCTGAAAGAGCTGGGA 2880
Db 22320 GCGCCCGAGGAAGAGCCAGGCTGCTTCCTTCCGGCTCTCAGGTCTGAAAGAGCTGGGA 22379
Qy 2881 CCGAGCGCGAGATTCCTTAAGGAGCTCTCAGATGAGACTGTGCTTGGGCGAGTCAGT 2940
Db 22380 CCGAGCGCGAGATTCCTTAAGGAGCTCTCAGATGAGACTGTGCTTGGGCGAGTCAGT 22439
Qy 2941 GACACTGGCTGCCAGGTGTAGCCAGCAGCTGCGCAGGCGACCTGGAGCAAGAGCG 3000
Db 22440 GACACTGGCTGCCAGGTGTAGCCAGCAGCTGCGCAGGCGACCTGGAGCAAGAGCG 22499
Qy 3001 A 3001
Db 22500 A 22500

RESULT 5

US-10-077-130-4
; Sequence 4, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MP12001-047P1RCP1(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 24120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(71)
; NAME/KEY: CDS
; LOCATION: (72)...(23978)
; NAME/KEY: 3'UTR
; LOCATION: (23979)...(24120)
US-10-077-130-4
Query Match 100.0%; Score 3001; DB 13; Length 24120;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 CCTACGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTGGCGCGCTGAG 60
Db 19571 CCTACGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTGGCGCGCTGAG 19630
Qy 61 CCACCCCTGTGTACGGGGCTGTGGACCAAGTTTGAGACCGGCAAGACCTCATCTCAT 120
Db 19631 CCAACCGCTGTGTACGGGGCTGTGGACCAAGTTTGAGACCGGCAAGACCTCATCTCAT 19690
Qy 121 CCTGGAGCTGTGTATCTCCAGGAGCTGTGGACCGCTGTACAGGAAGGCGCTGTGAC 180
Db 19691 CCTGGAGCTGTGTATCTCCAGGAGCTGTGGACCGCTGTACAGGAAGGCGCTGTGAC 19750
Qy 181 GGAGCGCAGGTCAAGGTCTACATCCAGAGCTGTGGAGGGGTGCACTACCTGCACAG 240
Db 19751 GGAGCGCAGGTCAAGGTCTACATCCAGAGCTGTGGAGGGGTGCACTACCTGCACAG 19810

Qy 241 CCATGGGGTCTTCCACCTGGACATAAAGCCCTCTAAACATCTGATGTTGTCATCTCTGCCG 300
Db 19811 CCATGGGGTCTTCCACCTGGACATAAAGCCCTCTAAACATCTGATGTTGTCATCTCTGCCG 19870
Qy 301 GGAAGACATTAANAATCTGGGACTTTGGCTTTGGCCGAAACATCAACCCAGCAGAGCTGCA 360
Db 19871 GGAAGACATTAANAATCTGGGACTTTGGCTTTGGCCGAAACATCAACCCAGCAGAGCTGCA 19930
Qy 361 GTTCAGCCAGTACCGGCTCCCTGAGTTGCTCTCCCGAGATCATCTCCAGCAGAACCTGT 420
Db 19931 GTTCAGCCAGTACCGGCTCCCTGAGTTGCTCTCCCGAGATCATCTCCAGCAGAACCTGT 19990
Qy 421 GAGGAAGCCCTCCGACATTTGGGCGCATGGGTGTCTCTCTACCTCAGCCCTGACTGCTC 480
Db 19991 GAGGAAGCCCTCCGACATTTGGGCGCATGGGTGTCTCTCTACCTCAGCCCTGACTGCTC 20050
Qy 481 ATCCCAATTTGCGGCGAGAGTGACCGTGCACCTCTCTGAAACGTCTCTGAGGGGCGCGT 540
Db 20051 ATCCCAATTTGCGGCGAGAGTGACCGTGCACCTCTCTGAAACGTCTCTGAGGGGCGCGT 20110
Qy 541 GTCATGAGCAGCCCATGCTGCCACCTCAGGAAGAGCGCAAGACTTCTCAAGGC 600
Db 20111 GTCATGAGCAGCCCATGCTGCCACCTCAGGAAGAGCGCAAGACTTCTCAAGGC 20170
Qy 601 TAGCTGACAGAGCCCTCAGGCGCGGCTAGTGGCGCCAGTGCCTCTCCACACCCCTG 660
Db 20171 TAGCTGACAGAGCCCTCAGGCGCGGCTAGTGGCGCCAGTGCCTCTCCACACCCCTG 20230
Qy 661 GTTCTGAAATTCATGCTCGGAGGAGGCGCATCTCATCAACCAAGCAGCTCAAGTT 720
Db 20231 GTTCTGAAATTCATGCTCGGAGGAGGCGCATCTCATCAACCAAGCAGCTCAAGTT 20290
Qy 721 CTTCTGGCCCGAAAGTCTGCGCAGCGTTCCCTGATGAGCTCAAGTCCATCTCTGTGAT 780
Db 20291 CTTCTGGCCCGAAAGTCTGCGCAGCGTTCCCTGATGAGCTCAAGTCCATCTCTGTGAT 20350
Qy 781 GCGCTCCATCCTGAGTGTGCTGGGGGCCACCCGACAGCCCTCTCTGCGCGTAGCCCG 840
Db 20351 GCGCTCCATCCTGAGTGTGCTGGGGGCCACCCGACAGCCCTCTCTGCGCGTAGCCCG 20410
Qy 841 GCACCTCTGACAGGACACTGTTGGCTCTCTCAGTTCTCTCTCTCTCTGACAAACGAGCT 900
Db 20411 GCACCTCTGACAGGACACTGTTGGCTCTCTCAGTTCTCTCTCTCTCTGACAAACGAGCT 20470
Qy 901 CGCCCAATTTGCCGGGCTAAGTCACTGCCACCTCCCGGTGACACACTCACACTGCT 960
Db 20471 CGCCCAATTTGCCGGGCTAAGTCACTGCCACCTCCCGGTGACACACTCACACTGCT 20530
Qy 961 GCAACCCCGGGGCTTCTGCGGCTCTGCGCCAGCTGCTGAGGAAGCCGAGCCAGTGA 1020
Db 20531 GCAACCCCGGGGCTTCTGCGGCTCTGCGCCAGCTGCTGAGGAAGCCGAGCCAGTGA 20590
Qy 1021 GCGCTCCACGAGGCGCCAGCTCCGCTGCACTCTCCGAGGTTGCGGGCCACCGGCGCG 1080
Db 20591 GCGCTCCACGAGGCGCCAGCTCCGCTGCACTCTCCGAGGTTGCGGGCCACCGGCGCG 20650
Qy 1081 CCAGGCTCGTGTGCGCCCGGCGACAGCTCATCGCAGCTGTTCTACACACAGCGGGTGA 1140
Db 20651 CCAGGCTCGTGTGCGCCCGGCGACAGCTCATCGCAGCTGTTCTACACACAGCGGGTGA 20710
Qy 1141 GAGCCCTGAGCAGCGGGGCTTCTGCGGCTCTGCGCCAGCTGCTGAGGAAGCCGCGCGCA 1200
Db 20711 GAGCCCTGAGCAGCGGGGCTTCTGCGGCTCTGCGCCAGCTGCTGAGGAAGCCGCGCGCA 20770
Qy 1201 CTTGCTGAAGGCGGGTACATTTCCGGGGCGCTGCCAGGCTGCGGAGCCACTGATGA 1260
Db 20771 CTTGCTGAAGGCGGGTACATTTCCGGGGCGCTGCCAGGCTGCGGAGCCACTGATGA 20830
Qy 1261 GCAACCGCTGTGTGAGGAGGAGCGCCAGGAGGAGGAGCCAGCTCTCTGGCCAAAGC 1320
Db 20831 GCAACCGCTGTGTGAGGAGGAGGCGCCAGGAGGAGGAGCCAGCTCTCTGGCCAAAGC 20890
Qy 1321 CCCTCATTCGAGACTGCGCTCGGCTGCTGCGCACCCACTTGGGCCCTGGGCCA 1380

[illegible]

Db	21971	CGAGTGGACCCCGCCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACTCCCATTT	220330
Qy	2461	CGAGTTTATGATCTTTCAGAAAGTCCCCAAGTCCGCTCAGCCAGAGCGCGCCTCCCCCAT	2520
Db	22031	CGAGTTTATGATCTTTCAGAAAGTCCCCAAGTCCGCTCAGCCAGAGCGCGCCTCCCCCAT	22090
Qy	2521	GGCTGAGGAGAGCTGGCGGAGTTCCGAGAGCCACGTCGCCCTTGCCAGGTGAACCTGGG	2580
Db	22091	GGCTGAGGAGAGCTGGCGGAGTTCCGAGAGCCACGTCGCCCTTGCCAGGTGAACCTGGG	22150
Qy	2581	CCCCCAGCCAGCGCTGGAGATCACAGAGAGTCAAGAGTGTGGAAGCGCTGCTGGGCAGA	2640
Db	22151	CCCCCAGCCAGCGCTGGAGATCACAGAGAGTCAAGAGTGTGGAAGCGCTGCTGGGCAGA	22210
Qy	2641	GGCTGCCGTGGGCAGGAAGCGAAGTGGTCTCGCCGTCACGCAGCGCTCTTTCACCTTCCC	2700
Db	22211	GGCTGCCGTGGGCAGGAAGCGAAGTGGTCTCGCCGTCACGCAGCGCTCTTTCACCTTCCC	22270
Qy	2701	TGGGAGGCACCTGCCCGCTGGATGAGCCTGCAGAGCTGGGGCTGCGTGAGAGAGTGAAGGC	2760
Db	22271	TGGGAGGCACCTGCCCGCTGGATGAGCCTGCAGAGCTGGGGCTGCGTGAGAGAGTGAAGGC	22330
Qy	2761	CTCCGTGGAGACATCTCCCGGATCTGAAGGGCAGGCCGGAAGGTCTGGAAGAAGGAGGG	2820
Db	22331	CTCCGTGGAGACATCTCCCGGATCTGAAGGGCAGGCCGGAAGGTCTGGAAGAAGGAGGG	22390
Qy	2821	GCCCCCAGGAAGACAGCGCTTGTCTTCTTCCGCTCTCAGGTCTGAAGAGCTGGGA	2880
Db	22391	GCCCCCAGGAAGACAGCGCTTGTCTTCTTCCGCTCTCAGGTCTGAAGAGCTGGGA	22450
Qy	2881	CCGAGCCCGACATTTCTAAGGAGCTCTCAGATGAGACTGTGGTCTCTGGGCCAGTCAGT	2940
Db	22451	CCGAGCCCGACATTTCTAAGGAGCTCTCAGATGAGACTGTGGTCTCTGGGCCAGTCAGT	22510
Qy	2941	GACACTGGCCTGCCAGGTGTCAAGCCAGCCAGCTGCCAGGCCACCTTGGAGCAAGACGG	3000
Db	22511	GACACTGGCCTGCCAGGTGTCAAGCCAGCCAGCTGCCAGGCCACCTTGGAGCAAGACGG	22570
Qy	3001	A 3001	
Db	22571	A 22571	

RESULT 6

```

RESULI 6
US-10-307-019-3
; Sequence 3, Application US/10307019
; Publication No. US20030108533A1
; GENERAL INFORMATION:
; APPLICANT: Zeng, Wenlin
; APPLICANT: Stanton, Lawrence
; APPLICANT: SCIOS, INC.
; TITLE OF INVENTION: MEDIATORS OF SIGNAL TRANSDUCTION
; FILE REFERENCE: SCIOS.0210V1
; CURRENT APPLICATION NUMBER: US/10/307,019
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/548,473
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/129,552
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5007
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (97) ... (4926)
US-10-307-019-3

```

Query Match	99.9%	Score 2999.4;	DB 15;	Length 5007;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3000;	Conservative	0;	Mismatches	1;
			Indels	0;
			Gaps	0;

QY 1 CCTACGAGCAGAACTCGGGGCCAGGCATACAGGAGCGAGACATCTCTGGCGGCGTGG 60
DB 522 CCTACGAGCAGAACTCGGGGCCAGGCATACAGGAGCGAGACATCTCTGGCGGCGTGG 581
QY 61 CCACCCGCTGCTACGGGGCTGCTGGACCAAGTTTGAGACCCGCAAGACCTCATCTTCAT 120
DB 582 CCACCCGCTGCTACGGGGCTGCTGGACCAAGTTTGAGACCCGCAAGACCTCATCTTCAT 641
QY 121 CTTGAGCTGTGCTCATCCGAGGAGCTCTGGACCGCTGTACAGGAAGGCGCTGTGAC 180
DB 642 CTTGAGCTGTGCTCATCCGAGGAGCTCTGGACCGCTGTACAGGAAGGCGCTGTGAC 701
QY 181 GGAGCCCGAGGTCGAAGGTCATACACAGCTGTGTGAGGGGCTGCACTACCTGCAAG 240
DB 702 GGAGCCCGAGGTCGAAGGTCATACACAGCTGTGTGAGGGGCTGCACTACCTGCAAG 761
QY 241 CCATGGGCTTCTCCACCTGGACATAAAGCCCTCTAAACATCTGTATGTGATCTGTGCCG 300
DB 762 CCATGGGCTTCTCCACCTGGACATAAAGCCCTCTAAACATCTGTATGTGATCTGTGCCG 821
QY 301 GGAAGACATTAATAATCTCGACATTTGGCTTTGGCCAGAACATCAACCCAGCAGAGCTGCA 360
DB 822 GGAAGACATTAATAATCTCGACATTTGGCTTTGGCCAGAACATCAACCCAGCAGAGCTGCA 881
QY 361 GTTCAGCAGTACGGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGCAGAAACCTGT 420
DB 882 GTTCAGCAGTACGGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGCAGAAACCTGT 941
QY 421 GAGCGAAGCTCCGACATTTGGCCATGGGTGTCATCTCTACCTGACCTGACCTGCTC 480
DB 942 GAGCGAAGCTCCGACATTTGGCCATGGGTGTCATCTCTACCTGACCTGACCTGCTC 1001
QY 481 ATCCCCATTTGCCCGGAGAGTGACCGCTGCCACCTCTCTGAACGCTCTGGAGGGCGCGT 540
DB 1002 ATCCCCATTTGCCCGGAGAGTGACCGCTGCCACCTCTCTGAACGCTCTGGAGGGCGCGT 1061
QY 541 GTCATGAGCAGCCCATGCTGCTGCCACCTCTGAGCGAAGCGCCAAAGACTTCATCAAGGC 600
DB 1062 GTCATGAGCAGCCCATGCTGCTGCCACCTCTGAGCGAAGCGCCAAAGACTTCATCAAGGC 1121
QY 601 TACGCTCAGAGACCCCTCAGGCGCGGCTAGTGGCGGCCAGTCCCTCTCCACCCCTG 660
DB 1122 TACGCTCAGAGACCCCTCAGGCGCGGCTAGTGGCGGCCAGTCCCTCTCCACCCCTG 1181
QY 661 GTTCCTGAATCCATGCTCGGAGGAGGCCATTCATCAACACCAAGCAGCTCAAGTT 720
DB 1182 GTTCCTGAATCCATGCTCGGAGGAGGCCATTCATCAACACCAAGCAGCTCAAGTT 1241
QY 721 CTTCTGGCCGAGTCTGCTGGCAGCGTTCTCTGATGAGCTACAAGTTCATCTCTGTGAT 780
DB 1242 CTTCTGGCCGAGTCTGCTGGCAGCGTTCTCTGATGAGCTACAAGTTCATCTCTGTGAT 1301
QY 781 GCGCTCCATCCCTGAGCTGCTGGGGGCCCAACCCGACAGCCCTCTCTGGCGGTAGCCCG 840
DB 1302 GCGCTCCATCCCTGAGCTGCTGGGGGCCCAACCCGACAGCCCTCTCTGGCGGTAGCCCG 1361
QY 841 GCACCTTGAGGAGACACTGGTGGCTCTCCAGTTCTCTCTCTCTCTGACAAACGAGCT 900
DB 1362 GCACCTTGAGGAGACACTGGTGGCTCTCCAGTTCTCTCTCTCTCTGACAAACGAGCT 1421
QY 901 CGCCCCATTTGCCGGGCTTAAGTCACTGGCAACCTCTCCCGGTGACACACTCACCTGCT 960
DB 1422 CGCCCCATTTGCCGGGCTTAAGTCACTGGCAACCTCTCCCGGTGACACACTCACCTGCT 1481
QY 961 GCACCCCGGGGCTTCTGCGGCCCTCGGCGAGCTCTGAGAGCCCTGAGAGCGAGGCGAGTGA 1020
DB 1482 GCACCCCGGGGCTTCTGCGGCCCTCGGCGAGCTCTGAGAGAGCGAGGCGAGTGA 1541
QY 1021 GCGCTCACGAGGCCCTCAGCTCCGCTGCACTCCCGAGGCTGCTCGGGCCACCGGCCG 1080
DB 1542 GCGCTCACGAGGCCCTCAGCTCCGCTGCACTCCCGAGGCTGCTCGGGCCACCGGCCG 1601

QY 1081 CCAGGGCTGGTGCCTCCGGCACAGCGTCACTCCGAGCTGTCTTACCAACAGGCGGCTGA 1140
DB 1602 CCAGGGCTGGTGCCTCCGGCACAGCGTCACTCCGAGCTGTCTTACCAACAGGCGGCTGA 1661
QY 1141 GAGCCCTGAGACACGGGGCTTGGCCCTCGGGGAGCAGGCGGACCCCGGCCCGGGCGGCA 1200
DB 1662 GAGCCCTGAGACACGGGGCTTGGCCCTCGGGGAGCAGGCGGACCCCGGCCCGGGCGGCA 1721
QY 1201 CTTGCTGAAGGGGGGTACATTTGGGGGGCGCTGCCAGGCTCTGGCGAGCACTGTATGA 1260
DB 1722 CTTGCTGAAGGGGGGTACATTTGGGGGGCGCTGCCAGGCTCTGGCGAGCACTGTATGA 1781
QY 1261 GCACCGCGTCTCTGAGGAGGAGGCGCGCCAGGAGGAGCAGGCCACCTCTCTGGCCAAAGC 1320
DB 1782 GCACCGCGTCTCTGAGGAGGAGGCGCGCCAGGAGGAGGAGGAGGCCACCTCTCTGGCCAAAGC 1841
QY 1321 CCCTCATTTGAGACTGCTCCCTCGGCTGCTCTCTGCGACCCCACTTTGGCCCTCTGGCCA 1380
DB 1842 CCCTCATTTGAGACTGCTCCCTCGGCTGCTCTCTGCGACCCCACTTTGGCCCTCTGGCCA 1901
QY 1381 CAGCCACTCCCTGAAACATGACTCTCCGAGCACCCCCCGGCCCTCTCTCGAGGCGCTGCGG 1440
DB 1902 CAGCCACTCCCTGAAACATGACTCTCCGAGCACCCCCCGGCCCTCTCTCGAGGCGCTGCGG 1961
QY 1441 TGAGGCACAGCACTGCTCTTCAAGCCCCCTCGGGGGGGCCCTTATCAGGGACATGGGGCA 1500
DB 1962 TGAGGCACAGCACTGCTCTTCAAGCCCCCTCGGGGGGGCCCTTATCAGGGACATGGGGCA 2021
QY 1501 CCCTCAGGGCTCCAGCAGCTTCATCTCACTGTGTGGCCACCCAGGCACTGTCTCAGCCAGA 1560
DB 2022 CCCTCAGGGCTCCAGCAGCTTCATCTCACTGTGTGGCCACCCAGGCACTGTCTCAGCCAGA 2081
QY 1561 GAGGCCATCCCGGACAGCCCTTGGGGGAGCCAGCCCTTCTGTGCCACCCCAAGCAGGG 1620
DB 2082 GAGGCCATCCCGGACAGCCCTTGGGGGAGCCAGCCCTTCTGTGCCACCCCAAGCAGGG 2141
QY 1621 TTCTGCCCCCAGGAGGGCTGACAGCCCCCAGCAGTGTGCCCAATGCCCTCTCTGGCTC 1680
DB 2142 TTCTGCCCCCAGGAGGGCTGACAGCCCCCAGCAGTGTGCCCAATGCCCTCTCTGGCTC 2201
QY 1681 CTTCCCTCAGAGACTTTCGAAAGAGCCCGCTTAGTACCTCAAGCCCTTCTTGGGACA 1740
DB 2202 CTTCCCTCAGAGACTTTCGAAAGAGCCCGCTTAGTACCTCAAGCCCTTCTTGGGACA 2261
QY 1741 GCCCAGGACACCCCTTGCCTGCCAAAGAACAGCCCGCTTAAAGTGGAGCC 1800
DB 2262 GCCCAGGACACCCCTTGCCTGCCAAAGAACAGCCCGCTTAAAGTGGAGCC 2321
QY 1801 TGGAGACATCTCTCTTCTGGGAGGCGCAAAACCCGGCCCTCGAGTTCCCAAGGTCAGC 1860
DB 2322 TGGAGACATCTCTCTTCTGGGAGGCGCAAAACCCGGCCCTCGAGTTCCCAAGGTCAGC 2381
QY 1861 CTTCCAGGCGAGCTCTTCCCAAGTACGCTCCCTCAGGGTGGGCTCTCTCCAGTGGGCA 1920
DB 2382 CTTCCAGGCGAGCTCTTCCCAAGTACGCTCCCTCAGGGTGGGCTCTCTCCAGTGGGCA 2441
QY 1921 AGAGCTGTGCGCCCTCTCTGGATCGGAGGCTTGGACCCAGAGGCTGAGGATCTGTCCGA 1980
DB 2442 AGAGCTGTGCGCCCTCTCTGGATCGGAGGCTTGGACCCAGAGGCTGAGGATCTGTCCGA 2501
QY 1981 CTTCCACACCACTTGCAGCGGCTCAGAAACAGGGGACCATGCGGCAAGTTCTCCCTGGG 2040
DB 2502 CTTCCACACCACTTGCAGCGGCTCAGAAACAGGGGACCATGCGGCAAGTTCTCCCTGGG 2561
QY 2041 TGGTCCGGGGCTACGAGGGCTGGCTGTATGCGACCTTTGCTTTGGTGGAGATGC 2100
DB 2562 TGGTCCGGGGCTACGAGGGCTGGCTGTATGCGACCTTTGCTTTGGTGGAGATGC 2621
QY 2101 AGGGGCACTGTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGTCCCAAGTC 2160
DB 2622 AGGGGCACTGTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGCTGTGTCCCAAGTC 2681
QY 2161 GGAGGAGGAGGAGGAGGAGGAGGCGCGCTGAGTCCAGTCCGAGGAGCAGCAGAGGC 2220

Db 6443 A 6443

RESULT 8

US-09-858-664A-1

; Sequence 1, Application US/09858664A

; Patent No. US20020072491A1

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui, et al.

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL000927-CIP

; CURRENT APPLICATION NUMBER: US/09/858,664A

; CURRENT FILING DATE: 2001-05-17

; PRIOR APPLICATION NUMBER: 09/711,134

; PRIOR FILING DATE: 2000-11-11

; NUMBER OF SEQ.ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ.ID NO 1

; LENGTH: 5207

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-858-664A-1

Query Match 99.9%; Score 2997.8; DB 9; Length 5207;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTACGAGCAGAACTCGGGCCAGGATACAGGAGGAGGAGACATCTTGGCGCGCTGAG 60

Db 699 CCTACGAGCAGAACTCGGGCCAGGATACAGGAGGAGGAGACATCTTGGCGCGCTGAG 758

Qy 61 CCACCCGCTGTCACGCGGCTGCTGGACCACTTTGAGACCGCGAAGACCTCATCTCTCAT 120

Db 759 CCACCCGCTGTCACGCGGCTGCTGGACCACTTTGAGACCGCGAAGACCTCATCTCTCAT 818

Qy 121 CCTGAGCTGTGTCATCCGAGGAGCTGCTGGACCGCTGTACAGGAGGCGTGGTGAC 180

Db 819 CCTGAGCTGTGTCATCCGAGGAGCTGCTGGACCGCTGTACAGGAGGCGTGGTGAC 878

Qy 181 GGAGCGGAGTCAAGTCTACATCCAGCAGCTGGTGGAGGCGTGCATCTACCTGCACAG 240

Db 879 GGAGCGGAGTCAAGTCTACATCCAGCAGCTGGTGGAGGCGTGCATCTACCTGCACAG 938

Qy 241 CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAAACATCTCTGATGGTGCATCTGCGCG 300

Db 939 CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAAACATCTCTGATGGTGCATCTGCGCG 998

Qy 301 GGAAGACATTAAATCTCGCACTTTGGCTTTGGCTTTGCCAGAAATACACCCAGCAGAGTGCA 360

Db 999 GGAAGACATTAAATCTCGCACTTTGGCTTTGGCTTTGCCAGAAATACACCCAGCAGAGTGCA 1058

Qy 361 GTTCAGCAGTACCGCTCCCTGATGTTGCTCTCCCGAGATCATCCAGCAGAACCTGT 420

Db 1059 GTTCAGCAGTACCGCTCCCTGATGTTGCTCTCCCGAGATCATCCAGCAGAACCTGT 1118

Qy 421 GAGCGAAGCCTCCGACATTTGGGCGCATGGGTGTCTCTCTACCTCAGCCTGACCTGCTC 480

Db 1119 GAGCGAAGCCTCCGACATTTGGGCGCATGGGTGTCTCTCTACCTCAGCCTGACCTGCTC 1178

Qy 481 ATCCCCATTTGCCGGCGAGTACCGGTGCCCTCTCTGAACTCTCTGGAGGCGCGGT 540

Db 1179 ATCCCCATTTGCCGGCGAGTACCGGTGCCCTCTCTGAACTCTCTGGAGGCGCGGT 1238

Qy 541 GTCATGAGCAGCCCCATGGCTGCCACCTCAGCGAAGACGCCAAAGACTTCATCAAGGC 600

Db 1239 GTCATGAGCAGCCCCATGGCTGCCACCTCAGCGAAGACGCCAAAGACTTCATCAAGGC 1298

Qy 601 TACGCTCAGAGACCCCTCAGGCGCGGCTAGTGGCGGCCAGTCTCTCCACCCCTG 660

Db 1299 TACGCTCAGAGACCCCTCAGGCGCGGCTAGTGGCGGCCAGTCTCTCCACCCCTG 1358

Qy 661 GTTCCTGAAATTCATGCTCGGAGGAGGCCACTTCATCAACACCAAGCAGCTCAAGTT 720

Db 1359 GTTCCTGAAATTCATGCTCGGAGGAGGCCACTTCATCAACACCAAGCAGCTCAAGTT 1418

Qy 721 CCTCTGCGCCGAAAGTCTGCGGAGCGTTCTCTGATGAGCTACAAGTCCATCTCTGGTGAT 780

Db 1419 CCTCTGCGCCGAAAGTCTGCGGAGCGTTCTCTGATGAGCTACAAGTCCATCTCTGGTGAT 1478

Qy 781 GCGCTCCATCTCTGAGCTGTGCGGGGCCACCCGACAGACCCCTCTCTCGCGTAGCCCG 840

Db 1479 GCGCTCCATCTCTGAGCTGTGCGGGGCCACCCGACAGACCCCTCTCTCGCGTAGCCCG 1538

Qy 841 GCACCTCTGACGAGGACACTGGTGGCTCTCTCAGTTCTCTCTCTCTCTGACACAGGACT 900

Db 1539 GCACCTCTGACGAGGACACTGGTGGCTCTCTCAGTTCTCTCTCTCTCTGACACAGGACT 1598

Qy 901 CGCCCATTTTCCCGGGCTAAGTCACTGCCACCTCTCCCGGTGACACACTCACCACCTGCT 960

Db 1599 CGCCCATTTTCCCGGGCTAAGTCACTGCCACCTCTCCCGGTGACACACTCACCACCTGCT 1658

Qy 961 GCACCCCGGGGCTTCTGCGGGCTCTCGGCGCTCTGAGGAGCCGAGGCGGAGTGA 1020

Db 1659 GCACCCCGGGGCTTCTGCGGGCTCTCGGCGCTCTGAGGAGCCGAGGCGGAGTGA 1718

Qy 1021 GCGCTCCACGAGGCCCTCAGCTCGGCTGCTCTCCGAGGCTCGCGGCCACCGCGCGC 1080

Db 1719 GCGCTCCACGAGGCCCTCAGCTCGGCTGCTCTCCGAGGCTCGCGGCCACCGCGCGC 1778

Qy 1081 CCAGGGCTGCTGCGCGGCGACAGCTCATCCGAGCTGTCTACACACAGGCGGGTGA 1140

Db 1779 CCAGGGCTGCTGCGCGGCGACAGCTCATCCGAGCTGTCTACACACAGGCGGGTGA 1838

Qy 1141 GAGCCCTGAGCAGCGGGCCCTTGGCGCGGAGCAGCGGCACTCCGCGCGCGCGGCA 1200

Db 1839 GAGCCCTGAGCAGCGGGCCCTTGGCGCGGAGCAGCGGCACTCCGCGCGCGCGGCA 1898

Qy 1201 CTGCTGAAGCGCGGTACATTTGCGGGGGCTGCGAGGCTGCGCGAGCCACTGATGGA 1260

Db 1899 CTGCTGAAGCGCGGTACATTTGCGGGGGCTGCGAGGCTGCGCGAGCCACTGATGGA 1958

Qy 1261 GCACCGGCTGCTGAGGAGGAGCGCGCGAGGAGCAGCGGCACTCTCTGGGCCAAAGC 1320

Db 1959 GCACCGGCTGCTGAGGAGGAGCGCGCGAGGAGCAGCGGCACTCTCTGGGCCAAAGC 2018

Qy 1321 CCCCTCATTTGAGACTGCTCCGCTGCTGCTGCGCACTTGGGCCCTCTGGCGCA 1380

Db 2019 CCCCTCATTTGAGACTGCTCCGCTGCTGCTGCGCACTTGGGCCCTCTGGCGCA 2078

Qy 1381 CAGCCACTCTCTGGAACATGACTCTCGAGCAGCCCCCGCGCTCTCTCGAGGCGCTGCGG 1440

Db 2079 CAGCCACTCTCTGGAACATGACTCTCGAGCAGCCCCCGCGCTCTCTCGAGGCGCTGCGG 2138

Qy 1441 TGAGGACACGACTGCTTACGCCCCCTCGGGGGGGCCCCCTATCAGGGACATGGGGCA 1500

Db 2139 TGAGGACACGACTGCTTACGCCCCCTCGGGGGGGCCCCCTATCAGGGACATGGGGCA 2198

Qy 1501 CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGTGGCCACCCAGGCACTGTCTCAGCCAGA 1560

Db 2199 CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGTGGCCACCCAGGCACTGTCTCAGCCAGA 2258

Qy 1561 GAGGCCATCTCCCGGACAGCCCTTGGGGGCGAGCCAGCCCTTTTCTGCGACCCCAAGCAGGG 1620

Db 2259 GAGGCCATCTCCCGGACAGCCCTTGGGGGCGAGCCAGCCCTTTTCTGCGACCCCAAGCAGGG 2318

Qy 1621 TTCTGCCCCCGAGGAGGCTGACGCCCCCAAGCAGTGGCCCCCTCTCTCTCTCTCTCTCTCT 1680

Db 2319 TTCTGCCCCCGAGGAGGCTGACGCCCCCAAGCAGTGGCCCCCTCTCTCTCTCTCTCTCTCT 2378

Qy 1681 CTTCCTCTCCAGGATCTTTGCAAAAGAGGCGCCCCCTTAGTACCTCTCAAGCCCTCTTTGGGACA 1740

Db 2379 CTTCCTCTCCAGGATCTTTGCAAAAGAGGCGCCCCCTTAGTACCTCTCAAGCCCTCTTTGGGACA 2438

QY 1741 GCCCAGGACCCCTGCCCCCTGCCAAGCAAGCAAGCCCCCATTTGGACTCTAAGATGGGCC 1800
Db 2439 GCCCAGGACCCCTGCCCCCTGCCAAGCAAGCAAGCCCCCATTTGGACTCTAAGATGGGCC 2498
QY 1801 TGGAGACATCTCTCTCTCTGGAGAGCCAAACCCCGGCCCTGCACTCTCCAGGGTCAGC 1860
Db 2499 TGGAGACATCTCTCTCTGGAGAGCCAAACCCCGGCCCTGCACTCTCCAGGGTCAGC 2558
QY 1861 CTCCAGGCGAGCTCTTCCCAAGTGAAGTCTCCCTCAGGGTGGGCTCTCTCCAGAGTGGGCAC 1920
Db 2559 CTCCAGGCGAGCTCTTCCCAAGTGAAGTCTCCCTCAGGGTGGGCTCTCTCCAGAGTGGGCAC 2618
QY 1921 AGAGCTGGCCCTCTCCCTGGATGCCGAGGGCTGGACCCAGAGGCTGAGATCTGTCCGA 1980
Db 2619 AGAGCTGGCCCTCTCCCTGGATGCCGAGGGCTGGACCCAGAGGCTGAGATCTGTCCGA 2678
QY 1981 CTCCACACCCACTTTGACGGGCTCAGGAACAGGCGACCATGCGCAAGTTCTCCCTGGG 2040
Db 2679 CTCCACACCCACTTTGACGGGCTCAGGAACAGGCGACCATGCGCAAGTTCTCCCTGGG 2738
QY 2041 TGGTCGGGGGCTACCGAGGCGTGGCTGGCTATGGCACCTTTGCTTTGGTGGAGATGC 2100
Db 2739 TGGTCGGGGGCTACCGAGGCGTGGCTGGCTATGGCACCTTTGCTTTGGTGGAGATGC 2798
QY 2101 AGGGGCTATGCTGGGCGAGGGCCCATGCTGGGCGAGGATAGCTGGGCTGTGTCCAGTC 2160
Db 2799 AGGGGCTATGCTGGGCGAGGGCCCATGCTGGGCGAGGATAGCTGGGCTGTGTCCAGTC 2858
QY 2161 GGAGGAGGAGGAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2220
Db 2859 GGAGGAGGAGGAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2918
QY 2221 CAGGCTGAGAGCCCATGCCCCAGGTCAGTGCAGGCGCTGTGCTGAGTCCGAGGCGC 2280
Db 2919 CAGGCTGAGAGCCCATGCCCCAGGTCAGTGCAGGCGCTGTGCTGAGTCCGAGGCGC 2978
QY 2281 TCCACAGGAGCTCTCCAGAGCCACCCCATGGAGGACATCGGGAGGTCTCCCTGGT 2340
Db 2979 TCCACAGGAGCTCTCCAGAGCCACCCCATGGAGGACATCGGGAGGTCTCCCTGGT 3038
QY 2341 GCAGATCCGGACCTGTGAGTGTGGGAGGCGGCGGACACAAATATCCCTGGACATTTTC 2400
Db 3039 GCAGATCCGGACCTGTGAGTGTGGGAGGCGGCGGACACAAATATCCCTGGACATTTTC 3098
QY 2401 CGAGGTGACCCCGCTACCTCAACCTCTCAGACCTGTACGATCAAGTACCTCCCAT 2460
Db 3099 CGAGGTGACCCCGCTACCTCAACCTCTCAGACCTGTACGATCAAGTACCTCCCAT 3158
QY 2461 CGAGTTTATGATCTTCAGGAAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCCAT 2520
Db 3159 CGAGTTTATGATCTTCAGGAAGTCCCAAGTCCGCTCAGCAGAGCGCCCTCCCCAT 3218
QY 2521 GGCTGAGGAGAGCTGGCCGAGTTCCCGAGGCCACGCTGGCCCTGGCAGGTGAATGGG 2580
Db 3219 GGCTGAGGAGAGCTGGCCGAGTTCCCGAGGCCACGCTGGCCCTGGCAGGTGAATGGG 3278
QY 2581 CCCCCAGCAGGCTGAGATCAAGAGAGTCAAGAGATGTGACGCGCTGTGCGAGA 2640
Db 3279 CCCCCAGCAGGCTGAGATCAAGAGAGTCAAGAGATGTGACGCGCTGTGCGAGA 3338
QY 2641 GGCTGCGGTGGGAGGAGCGCAAGTGTCTCGCGTCAAGGAGCTCTTCCACTTCCC 2700
Db 3339 GGCTGCGGTGGGAGGAGCGCAAGTGTCTCGCGTCAAGGAGCTCTTCCACTTCCC 3398
QY 2701 TGGAGGCACTGCGCTGATGAGCTGAGAGCTGGGCTGCGTGAAGTGAAGGC 2760
Db 3399 TGGAGGCACTGCGCTGATGAGCTGAGAGCTGGGCTGCGTGAAGTGAAGGC 3458
QY 2761 CTCGTTGAGCAGATCTCCCGGATCTTGAAGGCGAGCCCGGAAGTCTGGAAGAGG 2820
Db 3459 CTCGTTGAGCAGATCTCCCGGATCTTGAAGGCGAGCCCGGAAGTCTGGAAGAGG 3518
QY 2821 GCCCCCCAGGAAGAGCCAGGCTTGTCTTCTCCGCTCTCAGGTCTCTGAAGAGCTGGGA 2880

Db 3519 GCCCAGGAGGAAGCCAGGCTTGTCTTCCGGCTCTCAGGTCTGAAGAGCTGGGA 3578
QY 2881 CCAGAGCCCGACATCTCTAAGGGAGCTCTCAGATGACATGTGTCTCTGGGCCAGTCAGT 2940
Db 3579 CCAGAGCCCGACATCTCTAAGGGAGCTCTCAGATGACATGTGTCTCTGGGCCAGTCAGT 3638
QY 2941 GACACTGGCTGCGCAGGTGTGAGCCAGGCTGCGCCAGGCGACCTGGAGCAAGACGG 3000
Db 3639 GACACTGGCTGCGCAGGTGTGAGCCAGGCGACCTGGAGCAAGACGG 3698
QY 3001 A 3001
Db 3699 A 3699

RESULT 9
US-10-274-978-1
; Sequence 1, Application US/10274978
; Publication No. US20030064475A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274, 978
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ. ID NO. 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Human
US-10-274-978-1

Query Match 99.9%; Score 2997.8; DB 14; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CCTACGGAGCAGAACTCGGGCCCGAGGCATACAGGAGCGAGACATCCTGGCCGCGCTGAG 60
Db 699 CCTACGGAGCAGAACTCGGGCCCGAGGCATACAGGAGCGAGACATCCTGGCCGCGCTGAG 758
QY 61 CCACCCGCTGGTCA CGGGCTGTGAGCCAGTTTGAGACCCGCAAGACCTCATCTCAT 120
Db 759 CCACCCGCTGGTCA CGGGCTGTGAGCCAGTTTGAGACCCGCAAGACCTCATCTCAT 818
QY 121 CTTGGAGCTGTCTCATCCGAGAGCTGTGACCCGCTGTACAGGAAGGGGTGTGTGAC 180
Db 819 CTTGGAGCTGTCTCATCCGAGAGCTGTGACCCGCTGTGTACAGGAAGGGGTGTGTGAC 878
QY 181 GGAGGCGCAGGTCAGAGTCTACATCCAGCAGCTGTGTGAGGGGCTGCACCTACTGCAAG 240
Db 879 GGAGGCGCAGGTCAGAGTCTACATCCAGCAGCTGTGTGAGGGGCTGCACCTACTGCAAG 938
QY 241 CCATGGGCTTCTCCACCTGGACATTAAGCCCTCTAAATCTGTGATGTGTGATCTTCCCG 300
Db 939 CCATGGGCTTCTCCACCTGGACATTAAGCCCTCTAAATCTGTGATGTGTGATCTTCCCG 998
QY 301 GGAAGACATTAAATCTCGACCTTTGGCTTTTCCGCCAGAACATCAACCCAGCAGAGCTGCA 360
Db 999 GGAAGACATTAAATCTCGACCTTTGGCTTTTCCGCCAGAACATCAACCCAGCAGAGCTGCA 1058
QY 361 GTTCAGCAGTACGGCTCCCTGAGTTGCTTCCCGCAGATCATCCAGCAGAACCTTGT 420
Db 1059 GTTCAGCAGTACGGCTCCCTGAGTTGCTTCCCGCAGATCATCCAGCAGAACCTTGT 1118
QY 421 GAGCGAAGCTCCGACATTTGGGCCATTTGGGGCTCATCTTCTACCTCAGCTGACCTGCTC 480

QY 2641 GGCTGCGTGGGAGGAAGCGCAAGTGGTCTCGCGGTACGACAGCCTCTTCCACTTCCC 2700
DB 3339 GGCTGCGTGGGAGGAAGCGCAAGTGGTCTCGCGGTACGACAGCCTCTTCCACTTCCC 3398
QY 2701 TGGGAGGCACTTGCCTGCTGGATGAGCTGCGAGAGTGGGGTGTGCTGAGAGTGAAGGC 2760
DB 3399 TGGGAGGCACTTGCCTGCTGGATGAGCTGCGAGAGTGGGGTGTGCTGAGAGTGAAGGC 3458
QY 2761 CTCCGTGGAGCACATCTCCCGATCTGAAAGGCGAGCGGAGAGTCTGGAAGAGGG 2820
DB 3459 CTCCGTGGAGCACATCTCCCGATCTGAAAGGCGAGCGGAGAGTCTGGAAGAGGG 3518
QY 2821 GCCCCCAGGAAGAACCCAGGCTTGGCTTCTTCCCGGTCTCAGGTCTGAAAGAGCTGGGA 2880
DB 3519 GCCCCCAGGAAGAACCCAGGCTTGGCTTCTTCCCGGTCTCAGGTCTGAAAGAGCTGGGA 3578
QY 2881 CCGAGCGCGACATTCCTAAGGAGCTCTCAGATGAGACTGTGTGCTTGGGCCAGTCAGT 2940
DB 3579 CCGAGCGCGACATTCCTAAGGAGCTCTCAGATGAGACTGTGTGCTTGGGCCAGTCAGT 3638
QY 2941 GACACTGGCTGCCAGGTGTAGCCAGCAGCTGCCAGGCGCAGCTGCCAGGCGCAGTCAAG 3000
DB 3639 GACACTGGCTGCCAGGTGTAGCCAGCAGCTGCCAGGCGCAGTCAAG 3698
QY 3001 A 3001
DB 3699 A 3699
RESULT 10
US-10-274-978-3
; Sequence 3, Application US/10274978
; Publication No. US20030064475A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL00927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Human
US-10-274-978-3
Query Match 99.9%; Score 2997.8; DB 14; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CCTACGGAGCAGAACTCGGGCCAGGCATACAGGGAGCGAGACATCTCTGGCCGCGCTGAG 60
DB 699 CCTACGGAGCAGAACTCGGGCCAGGCATACAGGGAGCGAGACATCTCTGGCCGCGCTGAG 758
QY 61 CCACCCGCTGGTACGCGGGCTGTGGAACGATTTGAGACCCGCAAGACCTCTCAT 120
DB 759 CCACCCGCTGGTACGCGGGCTGTGGAACGATTTGAGACCCGCAAGACCTCTCAT 818
QY 121 CTTGAGCTGTGTCTATCCAGGAGCTGTGGAACGATTTGAGACCCGCAAGACCTCTCAT 180
DB 819 CTTGAGCTGTGTCTATCCAGGAGCTGTGGAACGATTTGAGACCCGCAAGACCTCTCAT 878
QY 181 GGAGGCGAGGTCAAGGTCTATCCAGCAGCTGTGGAAGGCGCTGCACTTACCTGCACAG 240
DB 879 GGAGGCGAGGTCAAGGTCTATCCAGCAGCTGTGGAAGGCGCTGCACTTACCTGCACAG 938

QY 241 CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAAACATCTCTGATGTGATCTGCCCCG 300
DB 939 CCATGGCGTTCTCCACCTGGACATAAAGCCCTCTAAACATCTCTGATGTGATCTGCCCCG 998
QY 301 GGAAGACATTTAAATCTGCGACTTTGGCTTTGCCAGAACATCACCCAGCAGAGCTGCA 360
DB 999 GGAAGACATTTAAATCTGCGACTTTGGCTTTGCCAGAACATCACCCAGCAGAGCTGCA 1058
QY 361 GTTCAGCAGTACGGCTCCCTGAGTTCCTCCCCGAGATCATCCAGAGAACCCCTGT 420
DB 1059 GTTCAGCAGTACGGCTCCCTGAGTTCCTCCCCGAGATCATCCAGAGAACCCCTGT 1118
QY 421 GAGGAGAGCTCCGACATTTGGGCGCATGGTGTCTCTACTCTAGCCCTGACTGCTC 480
DB 1119 GAGGAGAGCTCCGACATTTGGGCGCATGGTGTCTCTACTCTAGCCCTGACTGCTC 1178
QY 481 ATCCCAATTTGCCGGGAGAGTGACCGTGTGCACCTCTGAAAGTCTCTGGAGGGGCGCT 540
DB 1179 ATCCCAATTTGCCGGGAGAGTGACCGTGTGCACCTCTGAAAGTCTCTGGAGGGGCGCT 1238
QY 541 GTTCAGGAGAGCGCCCATGGCTGCCACCTCAGGAGAGCGCCAAAGACTTCATCAAGGC 600
DB 1239 GTTCAGGAGAGCGCCCATGGCTGCCACCTCAGGAGAGCGCCAAAGACTTCATCAAGGC 1298
QY 601 TAGCGTGCAGAGAGCCCTCAGGCGCGGCTAGTGGCGCCAGTGCCTCTCCACCCCTG 660
DB 1299 TAGCGTGCAGAGAGCCCTCAGGCGCGGCTAGTGGCGCCAGTGCCTCTCCACCCCTG 1358
QY 661 GTTCCTGAAATCCATGCTCGGAGAGGCGCCACTTTCATCAACACCAAGCAGCTCAAGTT 720
DB 1359 GTTCCTGAAATCCATGCTCGGAGAGGCGCCACTTTCATCAACACCAAGCAGCTCAAGTT 1418
QY 721 CTTCTGGCGCCGAGTTCGCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTCTGTGAT 780
DB 1419 CTTCTGGCGCCGAGTTCGCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTCTGTGAT 1478
QY 781 GCGCTCCATCTCCTGAGCTGCTGGGGCCACCCGACAGCCCTCTCGGGGTAGCCCCG 840
DB 1479 GCGCTCCATCTCCTGAGCTGCTGGGGCCACCCGACAGCCCTCTCGGGGTAGCCCCG 1538
QY 841 GCACCTCTGAGGAGACACTGGTGGTCTCTCCAGTTCTCTCTCTCTGTGACAAAGCT 900
DB 1539 GCACCTCTGAGGAGACACTGGTGGTCTCTCCAGTTCTCTCTCTCTGTGACAAAGCT 1598
QY 901 CGCCCATTTGGCGGCTTAAGTCACTGCGCACCTCTCCCGGTGACACACTCACCAGTGT 960
DB 1599 CGCCCATTTGGCGGCTTAAGTCACTGCGCACCTCTCCCGGTGACACACTCACCAGTGT 1658
QY 961 GCACCCCGGGGCTTCTGCGGCGCTCGGCGCAGCTGCTGAGGAGCGAGCGCAGTGA 1020
DB 1659 GCACCCCGGGGCTTCTGCGGCGCTCGGCGCAGCTGCTGAGGAGCGAGCGCAGTGA 1718
QY 1021 GCGCTCCACCGAGGCGCCAGCTGCGCTGCTCTCCAGGCTGCGGCGCACCAGCGCGC 1080
DB 1719 GCGCTCCACCGAGGCGCCAGCTGCGCTGCTCTCCAGGCTGCGGCGCACCAGCGCGC 1778
QY 1081 CCAGGCGCTGCGTGGCGCGCAGAGTCACTGCGCACCTGTTCTACCAACAGCGCGGTGA 1140
DB 1779 CCAGGCGCTGCGTGGCGCGCAGAGTCACTGCGCACCTGTTCTACCAACAGCGCGGTGA 1838
QY 1141 GAGCCCTGAGCACGCGGCGCTTGGCGCCCGGGAGCAGCGCGCACCCGCGCGCGCGCA 1200
DB 1839 GAGCCCTGAGCACGCGGCGCTTGGCGCCCGGGAGCAGCGCGCACCCGCGCGCGCGCA 1898
QY 1201 CTTGCTGAAGGCGCGGTACATTTGGGGGCGGTGCGAGGCTGCGCGAGGCACTCATGGA 1260
DB 1899 CTTGCTGAAGGCGCGGTACATTTGGGGGCGGTGCGAGGCTGCGCGAGGCACTCATGGA 1958
QY 1261 GCACCGCTGTGAGGAGGAGGCGCGCAGGAGGAGGAGGCGCACCTCTCGGCGCAAGC 1320
DB 1959 GCACCGCTGTGAGGAGGAGGCGCGCAGGAGGAGGAGGCGCACCTCTCGGCGCAAGC 2018


```
Qy 1321 CCCCTCATTGAGACTGCCCTCCGGCTGCCCTCTGGGACCCCACTTGGGCCCTTGGCCA 1380
Db 2019 CCCCTCATTCAGAGACTGCCCTCCGGCTGCCCTCTGGGACCCCACTTGGGCCCTTGGCCA 2078
Qy 1381 CAGCACTCCCTGGAACATGACTCTCCAGACACCCCGCCCTCTCCGAGGCTTGGG 1440
Db 2079 CAGCACTCCCTGGAACATGACTCTCCAGACACCCCGCCCTCTCCGAGGCTTGGG 2138
Qy 1441 TGAGGCAAGGCACTGGCTTCAGCCCCCTCCGGGGGGCCCCCTATCAAGGGACATGGGGCA 1500
Db 2139 TGAGGCAAGGCACTGGCTTCAGCCCCCTCCGGGGGGCCCCCTATCAAGGGACATGGGGCA 2198
Qy 1501 CCCTCAGGGCTCCAGAGACTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGA 1560
Db 2199 CCCTCAGGGCTCCAGAGACTTCCATCCACTGGTGGCCACCCAGGCACTGCTCAGCCAGA 2258
Qy 1561 GAGGCCATCCCGGACAGCCCTTGGGGGACCCAGCCCTTCTGCCACCCCAAGCAGGG 1620
Db 2259 GAGGCCATCCCGGACAGCCCTTGGGGGACCCAGCCCTTCTGCCACCCCAAGCAGGG 2318
Qy 1621 TTCTGCCCCCAGAGGGCTGAGCCCCCAACCCAGAGTTGCCCATGCCCTCTCTGGCTC 1680
Db 2319 TTCTGCCCCCAGAGGGCTGAGCCCCCAACCCAGAGTTGCCCATGCCCTCTCTGGCTC 2378
Qy 1681 CTTCCTCCAGGATCTTGCNAAGAGGCCCTTGTAGTACCTCAAGCCCTTCTTGGGACA 1740
Db 2379 CTTCCTCCAGGATCTTGCNAAGAGGCCCTTGTAGTACCTCAAGCCCTTCTTGGGACA 2438
Qy 1741 GCCCCAGGACCCCTGCCCCGCAAGAGCCGCCCATTTGGACTCTAAGATGGGGCC 1800
Db 2439 GCCCCAGGACCCCTGCCCCGCAAGAGCCGCCCATTTGGACTCTAAGATGGGGCC 2498
Qy 1801 TGGAGACATCTCTTCTCTGGAGGCCAAAACCCGGCCCCCTGCAGTTCTCCAGGGTCAGC 1860
Db 2499 TGGAGACATCTCTTCTCTGGAGGCCAAAACCCGGCCCCCTGCAGTTCTCCAGGGTCAGC 2558
Qy 1861 CTCCCAGGAGCTCTTCCCAAGTGAGCTCCCTCAGGTTGGCTCTCCAGAGTGGGCAC 1920
Db 2559 CTCCCAGGAGCTCTTCCCAAGTGAGCTCCCTCAGGTTGGCTCTCCAGAGTGGGCAC 2618
Qy 1921 AGAGCTGGCCCCCTCCCTGGATGGAGGGCTGACCCAGAGGCTGAGGATCTGTCCGA 1980
Db 2619 AGAGCTGGCCCCCTCCCTGGATGGAGGGCTGACCCAGAGGCTGAGGATCTGTCCGA 2678
Qy 1981 CTCACACCCCACTTGACGCGCCCTCAGGAACAGCGACCAATGCGCAAGTTCTCCCTGGG 2040
Db 2679 CTCACACCCCACTTGACGCGCCCTCAGGAACAGCGACCAATGCGCAAGTTCTCCCTGGG 2738
Qy 2041 TGGTCGGGGGGCTACGAGCGGTGGCTGGCTATGGCACTTGGCTTGGTGGAGATGC 2100
Db 2739 TGGTCGGGGGGCTACGAGCGGTGGCTGGCTATGGCACTTGGCTTGGTGGAGATGC 2798
Qy 2101 AGGGGGCATGTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGGCTGTGTCCAGTC 2160
Db 2799 AGGGGGCATGTGGGGCAGGGGCCCATGTGGGCCAGGATAGCTGGGGCTGTGTCCAGTC 2858
Qy 2161 GGAGGAGGAGGACGAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2220
Db 2859 GGAGGAGGAGGACGAGGAGGAGGCGCAGGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2918
Qy 2221 CAGGGCTGAGAGCCCATCTGCCCCAGGTCAAGTGCAGAGGCTGTGCTGAGGTCGCGAGGGC 2280
Db 2919 CAGGGCTGAGAGCCCATCTGCCCCAGGTCAAGTGCAGAGGCTGTGCTGAGGTCGCGAGGGC 2978
Qy 2281 TCCACACAGGAGCTCTCCAGAGCCACCCCATGGAGGACATCGGGCAGGTCCTCCCTGGT 2340
Db 2979 TCCACACAGGAGCTCTCCAGAGCCACCCCATGGAGGACATCGGGCAGGTCCTCCCTGGT 3038
Qy 2341 GCAGATCCGGGACCTGTGAGTGATGCGGAGGCGCGGACACAAATATCCCTGGACATTTTC 2400
Db 3039 GCAGATCCGGGACCTGTGAGTGATGCGGAGGCGCGGACACAAATATCCCTGGACATTTTC 3098
Qy 2401 CGAGGTGACCCCGCTACCTCAACCTCTCAGACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTT 2460
```

```
Db 3099 CGAGGTGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCATTT 3158
Qy 2461 CGAGTTTATGATCTTTCAGGAAAGTCCCAAGTCCCGCTCAGCAGAGCGCCCTCCCCCAT 2520
Db 3159 CGAGTTTATGATCTTTCAGGAAAGTCCCAAGTCCCGCTCAGCAGAGCGCCCTCCCCCAT 3218
Qy 2521 GGCTGAGGAGAGCTGGCCGAGTTCCCGAGGCCCACTGGGCCCTTGGCCAGTGAATGGG 2580
Db 3219 GGCTGAGGAGAGCTGGCCGAGTTCCCGAGGCCCACTGGGCCCTTGGCCAGTGAATGGG 3278
Qy 2581 CCCCCAGCAGGCTGAGATCACAGAGGAGTCAAGAGTGTGACGGCTGTCTGGCAGA 2640
Db 3279 CCCCCAGCAGGCTGAGATCACAGAGGAGTCAAGAGTGTGACGGCTGTCTGGCAGA 3338
Qy 2641 GGCTGCCGTGGGCAAGGAGCGCAAGTGTCTCCCGCTCAGCAGCCCTTTCCTTCCC 2700
Db 3339 GGCTGCCGTGGGCAAGGAGCGCAAGTGTCTCCCGCTCAGCAGCCCTTTCCTTCCC 3398
Qy 2701 TGGGAGCACCTGCCGCTGATGAGCTGAGAGCTGGGGCTGCTGAGAGAGTGAAGGC 2760
Db 3399 TGGGAGCACCTGCCGCTGATGAGCTGAGAGCTGGGGCTGCTGAGAGAGTGAAGGC 3458
Qy 2761 CTCGTCGAGCACATCTCCCGATCTTAAAGGGCAGGCCGGAAGTCTGGAGAAAGAGGG 2820
Db 3459 CTCGTCGAGCACATCTCCCGATCTTAAAGGGCAGGCCGGAAGTCTGGAGAAAGAGGG 3518
Qy 2821 GCCCCCAGGAAGAGCAGGCTTGTCTTCTCCGGCTCTCAGGTCTTGAAGAGCTGGGA 2880
Db 3519 GCCCCCAGGAAGAGCAGGCTTGTCTTCTCCGGCTCTCAGGTCTTGAAGAGCTGGGA 3578
Qy 2881 CCGAGCCCGACATCTTAAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCAGT 2940
Db 3579 CCGAGCCCGACATCTTAAAGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCAGT 3638
Qy 2941 GACACTGCCCTGCCAGGTGTCCAGCCAGCCAGCTGCCAGCCACCTGGAGCAAGAGCG 3000
Db 3639 GACACTGCCCTGCCAGGTGTCCAGCCAGCCAGCTGCCAGCCACCTGGAGCAAGAGCG 3698
Qy 3001 A 3001
Db 3699 A 3699
```

RESULT 11

```
US-10-697-263-1
; Sequence 1, Application US/10697263
; Publication No. US20040063142A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,263
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-697-263-1
```

```
Query Match 99.9%; Score 2997.8; DB 17; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```


QY 1 CTTACGAGCAGAACTCGGSCCAGGCATACAGGAGCGAGACATCCTTGCCGCGCTGAG 60
DB |||||
699 CTTACGAGCAGAACTCGGSCCAGGCATACAGGAGCGAGACATCCTTGCCGCGCTGAG 758
QY 61 CCACCCGCTGCTCACCGGSGCTGCTGACACAGTTTGAGACCCGAGAGACCCCTCATCTCAT 120
DB |||||
759 CCACCCGCTGCTCACCGGSGCTGCTGACACAGTTTGAGACCCGAGAGACCCCTCATCTCAT 818
QY 121 CTTGAGAGCTGTGCTCATCCGAGAGCTGTGACCGCCTGTACAGAGGGGCTGGTGAC 180
DB |||||
819 CTTGAGAGCTGTGCTCATCCGAGAGCTGTGACCGCCTGTACAGAGGGGCTGGTGAC 878
QY 181 GAGGCCGAGGCTCAAGGTTCTATCCAGCAGCTGTGAGGGGCTGCACTACTCTGCACAG 240
DB |||||
879 GAGGCCGAGGCTCAAGGTTCTATCCAGCAGCTGTGAGGGGCTGCACTACTCTGCACAG 938
QY 241 CCATGSCGTTTCCACCTCGACATTAAGGCCCTTAACATCTCTGATGTGCTTCCTGCCG 300
DB |||||
939 CCATGSCGTTTCCACCTCGACATTAAGGCCCTTAACATCTCTGATGTGCTTCCTGCCG 998
QY 301 GGAAGACATTAATAATCTGCGACTTTGGCTTTGCCAGAACATCAACCAGCAGAGCTGCA 360
DB |||||
999 GGAAGACATTAATAATCTGCGACTTTGGCTTTGCCAGAACATCAACCAGCAGAGCTGCA 1058
QY 361 GTTACGCCAGTACGGGCTCCCTGAGTTGCTCTCCGCCAGATCATCCAGCAGAACCCCTGT 420
DB |||||
1059 GTTACGCCAGTACGGGCTCCCTGAGTTGCTCTCCGCCAGATCATCCAGCAGAACCCCTGT 1118
QY 421 GAGCGAAGCCTCCGACATTTGGGCCATGSGGTGTATCTCTACCTCTGAGGAGGCGCGT 540
DB |||||
1179 ATCCCAATTTGCGCGCAGAGTGACCGGTGCCACCTCCTGAACTCTCTGAGGAGGCGCGT 1238
QY 541 GTCATGAGACACCCCATGCTGCCACCTCAGCGAAGACGCCAAGAATTTCATCAAGGC 600
DB |||||
1239 GTCATGAGACACCCCATGCTGCCACCTCAGCGAAGACGCCAAGAATTTCATCAAGGC 1298
QY 601 TACGCTGACAGAGCCCTCAGGCGCGGCTAGTGCGGCCAGTGCTCTCTCCACCCCTG 660
DB |||||
1299 TACGCTGACAGAGCCCTCAGGCGCGGCTAGTGCGGCCAGTGCTCTCTCCACCCCTG 1358
QY 661 GTTCTGAAATCCATGCTCGGAGGAGGCCCACTTCATCAACACCAAGCAGCTCAAGTT 720
DB |||||
1359 GTTCTGAAATCCATGCTCGGAGGAGGCCCACTTCATCAACACCAAGCAGCTCAAGTT 1418
QY 721 CCTCTGSCCGAAGTCTGTCGACAGCTTCCCTGATGAGCTACAAGTCCATCTCTGCTGAT 780
DB |||||
1419 CCTCTGSCCGAAGTCTGTCGACAGCTTCCCTGATGAGCTACAAGTCCATCTCTGCTGAT 1478
QY 781 GCGTCTCATCTCCTGAGCTGTGCGGGCCACACCCGACAGCCCTCTCCCTCGGCGTAGCCG 840
DB |||||
1479 GCGTCTCATCTCCTGAGCTGTGCGGGCCACACCCGACAGCCCTCTCCCTCGGCGTAGCCG 1538
QY 841 GCACCTCTGACGAGACACTGTGTGCTCCTCAGTTCTCTCTCTCTGACACAGCT 900
DB |||||
1539 GCACCTCTGACGAGACACTGTGTGCTCCTCAGTTCTCTCTCTCTGACACAGCT 1598
QY 901 CGCCCAATTTGCCCGGCTTAAGTCACTGACCACTCCCGGTGACACACTCACTGCT 960
DB |||||
1599 CGCCCAATTTGCCCGGCTTAAGTCACTGACCACTCCCGGTGACACACTCACTGCT 1658
QY 961 GCACCCCTCGGGCTTCTGCGGCCCTCGGCAGCTGCTGAGGAGGCGAGGCCAGTGCA 1020
DB |||||
1659 GCACCCCTCGGGCTTCTGCGGCCCTCGGCAGCTGCTGAGGAGGCGAGGCCAGTGCA 1718
QY 1021 GCGTCTCACCGAGGCCCACTCGCTGATCTCCGAGGCTGCGGGCCACCGGCCG 1080
DB |||||
1719 GCGTCTCACCGAGGCCCACTCGCTGATCTCCGAGGCTGCGGGCCACCGGCCG 1778

QY 1081 CCAGSGCTGCGTGCCTCCCGGCA CAGGCTATCCGCGAGCCTGTTCTTACACACAGGCGGGTGA 1140
DB |||||
1779 CCAGSGCTGCGTGCCTCCCGGCA CAGGCTATCCGCGAGCCTGTTCTTACACACAGGCGGGTGA 1838
QY 1141 GAGCCTCTGAGCA CCGGGCCCTTGGCCCCGGGAGCAGGCGGCA CCGCGCCCGGCGCGCGCA 1200
DB |||||
1839 GAGCCTCTGAGCA CCGGGCCCTTGGCCCCGGGAGCAGGCGGCA CCGCGCCCGGCGCGCGCA 1898
QY 1201 CTTGCTGAAGGGCGGGTACATTTGGGGGGCGCTGCGAGGCTGCGCAGCACTGATGGA 1260
DB |||||
1899 CTTGCTGAAGGGCGGGTACATTTGGGGGGCGCTGCCAGGCTGCGCAGCACTGATGGA 1958
QY 1261 GCACCGGCTGCTGAGGAGGAGCGCGCAGGAGGAGCAGGCGCA CCGCTCTGCGCCAAAGC 1320
DB |||||
1959 GCACCGGCTGCTGAGGAGGAGCGCGCAGGAGGAGCAGGCGCA CCGCTCTGCGCCAAAGC 2018
QY 1321 CCGCTCATTTGAGAGCTGCTCCGCTGCTGCTCTGCGCA CCGCATTTGCGCCCTCTGCGCA 1380
DB |||||
2019 CCGCTCATTTGAGAGCTGCTCCGCTGCTGCTCTGCGCA CCGCATTTGCGCCCTCTGCGCA 2078
QY 1381 CAGGCACTCCCTGGAACATGACTCTCGAGCA CCGCGCCCTCTCTCGAGGCGCTGCGG 1440
DB |||||
2079 CAGGCACTCCCTGGAACATGACTCTCGAGCA CCGCGCCCTCTCTCGAGGCGCTGCGG 2138
QY 1441 TGAGGCA CAGGACTGCTTTGAGCCCTCCGCGGGGCGCCCTATCAGGAGACATGSGGCA 1500
DB |||||
2139 TGAGGCA CAGGACTGCTTTGAGCCCTCCGCGGGGCGCCCTATCAGGAGACATGSGGCA 2198
QY 1501 CCTCTAGGGCTCCAGAGCTTCCATCCATGCTGCGCCACCCAGGCACTGCTCAGCCAGA 1560
DB |||||
2199 CCTCTAGGGCTCCAGAGCTTCCATCCATGCTGCGCCACCCAGGCACTGCTCAGCCAGA 2258
QY 1561 GAGGCCATCCCGGACAGCCCTTGGGGGAGCGAGCCCTTTCTGCGCA CCGCAAGCAGGG 1620
DB |||||
2259 GAGGCCATCCCGGACAGCCCTTGGGGGAGCGAGCCCTTTCTGCGCA CCGCAAGCAGGG 2318
QY 1621 TTCTGCCCCCAGGAGGCTGACGCCCCCA CCGAGCATTTGCCCCCA TGCCCTCTCTGGCTC 1680
DB |||||
2319 TTCTGCCCCCAGGAGGCTGACGCCCCCA CCGAGCATTTGCCCCCA TGCCCTCTCTGGCTC 2378
QY 1681 CTTTCCCTCAGAGATCTTGCAGAGCCCTTTAGTA CCGCTCAGGCCCTTTCTTGGGACA 1740
DB |||||
2379 CTTTCCCTCAGAGATCTTGCAGAGCCCTTTAGTA CCGCTCAGGCCCTTTCTTGGGACA 2438
QY 1741 GCGCCAGGCA CCGCCCTGCGCCCTGCCAAGCAAGCCCCCATTTGGA CTCTTAAGTGGGGC 1800
DB |||||
2439 GCGCCAGGCA CCGCCCTGCGCCCTGCCAAGCAAGCCCCCATTTGGA CTCTTAAGTGGGGC 2498
QY 1801 TGGAGACATCTCTCTTCTGAGGCGCAAAACCCCGGCCCTGCACTTCCCCAGGGTCAAGC 1860
DB |||||
2499 TGGAGACATCTCTCTTCTGAGGCGCAAAACCCCGGCCCTGCACTTCCCCAGGGTCAAGC 2558
QY 1861 CTTCCAGGCGAGCTCTTCCCAAGTGAAGTCTCTCAGGGTGGGCTCTCTCCAGGTGGGAC 1920
DB |||||
2559 CTTCCAGGCGAGCTCTTCCCAAGTGAAGTCTCTCAGGGTGGGCTCTCTCCAGGTGGGAC 2618
QY 1921 AGAGCTGCGGCCCTCTCTGGAATGCGAGGCTGGACCCAGAGGCTGAGGATCTCTCCGA 1980
DB |||||
2619 AGAGCTGCGGCCCTCTCTGGAATGCGAGGCTGGACCCAGAGGCTGAGGATCTCTCCGA 2678
QY 1981 CTCGACACCTCTTGCAGCGGCTCTCAGAACAGGCGCA CCGAGCACTTCTCCCTGGG 2040
DB |||||
2679 CTCGACACCTCTTGCAGCGGCTCTCAGAACAGGCGCA CCGAGCACTTCTCCCTGGG 2738
QY 2041 TGGTTCGCGGGGCTTACGAGCGCTGGCTGCTATGGCACTTTCCTTTGGTGGAGATGC 2100
DB |||||
2739 TGGTTCGCGGGGCTTACGAGCGCTGGCTGCTATGGCACTTTCCTTTTGGTGGAGATGC 2798
QY 2101 AGGGGCACTGCTGGGGCAGGGGCCATGTTGGGCCAGGATAGCCTGGGGTGTGTCCCACTC 2160
DB |||||
2799 AGGGGCACTGCTGGGGCAGGGGCCATGTTGGGGCAGGATAGCCTGGGGTGTGTCCCACTC 2858
QY 2161 GGAGGAGGAGGAGGAGGAGGCGCGGCTGATCTCCAGTCCGAGGAGGAGGAGGAGGCGC 2220

[illegible][illegible]

Db 3699 A 3699

RESULT 13

US-10-921-168-1

; Sequence 1, Application US/10921168

; Publication No. US20050003466A1

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui, et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: CL000927-CIP-DIV3

; CURRENT APPLICATION NUMBER: US/10/921.168

; CURRENT FILING DATE: 2004-08-19

; PRIOR FILING DATE: 10/274,978

; PRIOR FILING DATE: 2002-10-22

; PRIOR APPLICATION NUMBER: 09/858,664

; PRIOR FILING DATE: 2001-05-17

; PRIOR APPLICATION NUMBER: 09/711,134

; PRIOR FILING DATE: 2000-11-14

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1

; LENGTH: 5207

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-921-168-1

Query Match 99.9%; Score 2997.8; DB 18; Length 5207;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CCTACGGAGAGAACTCGGGCCCGAGGATACAGGAGGAGAGACATCTCTGGCGCGGCTGAG 60

Db 699 CCTACGGAGAGAACTCGGGCCCGAGGATACAGGAGGAGAGACATCTCTGGCGCGGCTGAG 758

Qy 61 CCACCCGCTGTACGGGGCTGTGGACCAAGTTTGAGACCCGCAAGACCTCATCTCTCAT 120

Db 759 CCACCCGCTGTGTACGGGGCTGTGGACCAAGTTTGAGACCCGCAAGACCTCATCTCTCAT 818

Qy 121 CCTGGAGCTGTCTATCCGAGGAGCTCTGGACCGCTGTACAGGAAGGCGGTGTGAC 180

Db 819 CCTGGAGCTGTCTATCCGAGGAGCTCTGGACCGCTGTACAGGAAGGCGGTGTGAC 878

Qy 181 GGAGCCGAGGTCAAGGTCTATCATCCAGCAGCTGTGGAGGGGCTGCACTACCTGCACAG 240

Db 879 GGAGCCGAGGTCAAGGTCTATCATCCAGCAGCTGTGGAGGGGCTGCACTACCTGCACAG 938

Qy 241 CCATGGCGTTCTCACCTGGACATAAAGCCCTCTAAACATCTGTATGTGTGATCTCTGGCCG 300

Db 939 CCATGGCGTTCTCACCTGGACATAAAGCCCTCTAAACATCTGTATGTGTGATCTCTGGCCG 998

Qy 301 GGAAGACATTAATCTGGACTTTGGCTTTGGCCAGAACATCAACCCAGAGAGCTGCA 360

Db 999 GGAAGACATTAATCTGGACTTTGGCTTTGGCCAGAACATCAACCCAGAGAGCTGCA 1058

Qy 361 GTTCAGCCAGTACGGCTCCCTGAGTTGCTCTCCCGGAGATCATCCAGCAGAACCTGT 420

Db 1059 GTTCAGCCAGTACGGCTCCCTGAGTTGCTCTCCCGGAGATCATCCAGCAGAACCTGT 1118

Qy 421 GAGCGAAGCCTCCGACATTTGGGCGATGGGTGTGATCTCTCACTCAAGCTGTGCTC 480

Db 1119 GAGCGAAGCCTCCGACATTTGGGCGATGGGTGTGATCTCTCACTCAAGCTGTGCTC 1178

Qy 481 ATCCCCATTTCCCGGAGAGTACCGCTGACACCTCTGACCTCTCTGGAGGGGCGGT 540

Db 1179 ATCCCCATTTCCCGGAGAGTACCGCTGACACCTCTCTGACCTCTCTGGAGGGGCGGT 1238

Qy 541 GTCATGAGCAGCCCACTGGCTGCGCCACTCAGCGAAGACGCAAGACTTTTCATCAAGGC 600

Db 1239 GTCATGAGCAGCCCACTGGCTGCGCCACTCAGCGAAGACGCAAGACTTTTCATCAAGGC 1298

Qy 601 TAGCTGCAGAGAGCCCTCAGGCCCGCGCTAGTGGCGGCCAGTGCCTCTCTCCACCCCTG 660

Db 1299 TAGCTGCAGAGAGCCCTCAGGCCCGCGCTAGTGGCGGCCAGTGCCTCTCTCCACCCCTG 1358

Qy 661 GTTCTGAAATCCATGCTCGGAGGAGGCCACTTATCAACACCAAGAGAGCTCAAGTT 720

Db 1359 GTTCTGAAATCCATGCTCGGAGGAGGCCACTTATCAACACCAAGAGAGCTCAAGTT 1418

Qy 721 CCTCTGGCCCGAAGTCTGTGGGAGCGTTCCCTGATGAGCTTACAGTCCATCTCTGGTGTAT 780

Db 1419 CCTCTGGCCCGAAGTCTGTGGGAGCGTTCCCTGATGAGCTTACAGTCCATCTCTGGTGTAT 1478

Qy 781 GCGCTCCATCCCTGAGTGTCTGGGGGCCACACCGAGAGCCCTCTCTGGCGGTAGCCCG 840

Db 1479 GCGCTCCATCCCTGAGTGTCTGGGGGCCACACCGAGAGCCCTCTCTGGCGGTAGCCCG 1538

Qy 841 GCACCTCTGCAGGGACACTGTGTGCTCTCCAGTTCTCTCTCTCTGACACAGAGCT 900

Db 1539 GCACCTCTGCAGGGACACTGTGTGCTCTCCAGTTCTCTCTCTCTGACACAGAGCT 1598

Qy 901 CGCCCAATTTGCCCGGGCTAAGTCACTGCACACCTCCCGGTGACACACTCACCAGCT 960

Db 1599 CGCCCAATTTGCCCGGGCTAAGTCACTGCACACCTCCCGGTGACACACTCACCAGCT 1658

Qy 961 GCACCCCGGGGCTTCTGTGGGCCCTCGGCAGCTCTGAGGAGCCGAGGCGCAGTGA 1020

Db 1659 GCACCCCGGGGCTTCTGTGGGCCCTCGGCAGCTCTGAGGAGCCGAGGCGCAGTGA 1718

Qy 1021 GCGTCTCAACGAGGCCCTCAGCTCCGCTGCTATCTCCGAGGGTCCGCGGCCACCGCGCCG 1080

Db 1719 GCGTCTCAACGAGGCCCTCAGCTCCGCTGCTATCTCCGAGGGTCCGCGGCCACCGCGCCG 1778

Qy 1081 CCAGGGTGTGTGCTCCCGGACAGCGTCACTCCGAGCTGTCTTACCAACAGGCGGGTGA 1140

Db 1779 CCAGGGTGTGTGCTCCCGGACAGCGTCACTCCGAGCTGTCTTACCAACAGGCGGGTGA 1838

Qy 1141 GAGCCCTGACACGGGGCCCTTGGCCCGGGGAGCAGGGGACACCGCCCGCGCGCGGCA 1200

Db 1839 GAGCCCTGACACGGGGCCCTTGGCCCGGGGAGCAGGGGACACCGCCCGCGCGCGGCA 1898

Qy 1201 CCTCTGAAGGGCGGTACTATTTGGGGGGCTCTCCAGGCTCTGCGAGCCCTGCGAGCCACTGATGGA 1260

Db 1899 CCTCTGAAGGGCGGTACTATTTGGGGGGCTCTCCAGGCTCTGCGAGCCACTGATGGA 1958

Qy 1261 GCACCGGCTGTGAGAGAGAGCGCCAGGAGGAGCAGGCGCACCTCTCTGGCCAAAGC 1320

Db 1959 GCACCGGCTGTGAGAGAGAGCGCCAGGAGGAGCAGGCGCACCTCTCTGGCCAAAGC 2018

Qy 1321 CCCCTGATTCGAGACTGCCCTCCGGCTGCTGCTCTGGCACCCACTTGGGCCCTTGCGCA 1380

Db 2019 CCCCTGATTCGAGACTGCCCTCCGGCTGCTGCTCTGGCACCCACTTGGGCCCTTGCGCA 2078

Qy 1381 CAGCCACTCTCTGGAACAATGACTCTCCGAGCACCCCGCCCTCTCTCGAGGGCTGCGG 1440

Db 2079 CAGCCACTCTCTGGAACAATGACTCTCCGAGCACCCCGCCCTCTCTCGAGGGCTGCGG 2138

Qy 1441 TGAGGCAACAGCACTGCTTACGCCCTCCGGGGGGCCCTATCAAGGACATGGGGCA 1500

Db 2139 TGAGGCAACAGCACTGCTTACGCCCTCCGGGGGGCCCTATCAAGGACATGGGGCA 2198

Qy 1501 CCCTCAGGGCTCCAGAGAGTTCATCCATCTGTGGCCACCCAGGCACTGTCTCAGCCAGA 1560

Db 2199 CCCTCAGGGCTCCAGAGAGTTCATCCATCTGTGTGGCCACCCAGGCACTGTCTCAGCCAGA 2258

Qy 1561 GAGGCCATCCCGGAGAGCCCTTGGGGGAGCAGCAGCCCTTTCTTGCCACCCCAAGAGAGG 1620

Db 2259 GAGGCCATCCCGGAGAGCCCTTGGGGGAGCAGCAGCCCTTTCTTGCCACCCCAAGAGAGG 2318

Qy 1621 TTCTGCCCGCCAGAGGGCTGACGCCCCCAACCCAGAGAGTTGCCCATCTCTCTGGCTC 1680

Db 2319 TTCTGCCCGCCAGAGGGCTGACGCCCCCAACCCAGAGAGTTGCCCATCTCTCTGGCTC 2378

Qy 1681 CTTCCCTCAGGATCTTGGCAAGAGGGCCCTTAGTACCTCTAAGGCCCTCTCTTGGGACA 1740

Db 2379 CTTCCCTCCAGGATCTTGAAGAGGCCCTTCTAGTACCTCAAGCCCTTCTTGGACA 2438
QY 1741 GCCCCCAGGACCCCTTCCCTTCCCAAAGCAAGCCCCCATTTGGAATCTTAAGATGGGCC 1800
Db 2439 GCCCCCAGGACCCCTTCCCTTCCCAAAGCAAGCCCCCATTTGGAATCTTAAGATGGGCC 2498
QY 1801 TGGAGACATCTCTTTCTTCTGGAGGCCAAAACCCGCCCCCTTCAGATTTCCCCAGGGTCAGC 1860
Db 2499 TGGAGACATCTCTTTCTTCTGGAGGCCAAAACCCGCCCCCTTCAGATTTCCCCAGGGTCAGC 2558
QY 1861 CTCCTCCAGGCGAGCTCTTCCCAAGTGAGCTCCCTCAGGCTGGGCTCCTCCAGGTGGGAC 1920
Db 2559 CTCCTCCAGGCGAGCTCTTCCCAAGTGAGCTCCCTCAGGCTGGGCTCCTCCAGGTGGGAC 2618
QY 1921 AGAGCCTGGCCCCCTCCCTCGATGCGGAGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 1980
Db 2619 AGAGCCTGGCCCCCTCCCTCGATGCGGAGGCTGGACCCAGGAGGCTGAGGATCTGTCCGA 2678
QY 1981 CTCACACCCACTTTCAGCGGCTTCAGAAACAGCGACCAATGCGCAAGTTCTCCCTGGG 2040
Db 2679 CTCACACCCACTTTCAGCGGCTTCAGAAACAGCGACCAATGCGCAAGTTCTCCCTGGG 2738
QY 2041 TGGTCCGCGGGCTACGACAGGCGTGGCTGGCTATGGCACTTGGCTTTGGTGGAGATGC 2100
Db 2739 TGGTCCGCGGGCTACGACAGGCGTGGCTGGCTATGGCACTTGGCTTTGGTGGAGATGC 2798
QY 2101 AGGGGGCATGTGGGGCAGGGGCCCATGTGGCCAGGATAGCTGGGCTGTGTCCAGTC 2160
Db 2799 AGGGGGCATGTGGGGCAGGGGCCCATGTGGCCAGGATAGCTGGGCTGTGTCCAGTC 2858
QY 2161 GGAGAGAGAGGACAGGAGGAGGCGCAGGGCTGAGTCCAGTGGAGGACGACGAGGAGC 2220
Db 2859 GGAGAGAGAGGACAGGAGGAGGCGCAGGGCTGAGTCCAGTGGAGGACGACGAGGAGC 2918
QY 2221 CAGGGCTCAGAGCCCACTGCCCACTCAGTGCAGGCGCTGCTGCTGAGTTCGCGAGGCG 2280
Db 2919 CAGGGCTCAGAGCCCACTGCCCACTCAGTGCAGGCGCTGCTGCTGAGTTCGCGAGGCG 2978
QY 2281 TCCCAACAGGAGCTCTCCAGAGCCCACTCCATGGAGGACATCGGGCAGGTCTCCCTGGT 2340
Db 2979 TCCCAACAGGAGCTCTCCAGAGCCCACTCCATGGAGGACATCGGGCAGGTCTCCCTGGT 3038
QY 2341 GCAGATCCGGACCTGTCCAGGTGATGCGGAGCGCCCGACACAATATCCCTGGACATTC 2400
Db 3039 GCAGATCCGGACCTGTCCAGGTGATGCGGAGCGCCCGACACAATATCCCTGGACATTC 3098
QY 2401 CGAGTGGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
Db 3099 CGAGTGGACCCCGCTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 3158
QY 2461 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCGAGCGCCCTCCCCCAT 2520
Db 3159 CGAGTTTATGATCTTCAGGAAAGTCCCAAGTCCGCTCAGCGAGCGCCCTCCCCCAT 3218
QY 2521 GGCTGAGGAGAGCTGGCGAGTTCCCGAGGCCCGCTGGCGCTTGGCGAGTGAATGGG 2580
Db 3219 GGCTGAGGAGAGCTGGCGAGTTCCCGAGGCCCGCTGGCGCTTGGCGAGTGAATGGG 3278
QY 2581 CCCCACGCGAGGCTGGAGATCAAGAGGAGTCAAGAGATGTGACCGCGCTGTGGCAGA 2640
Db 3279 CCCCACGCGAGGCTGGAGATCAAGAGGAGTCAAGAGATGTGACCGCGCTGTGGCAGA 3338
QY 2641 GGCTGCCGTGGCAGGAAGCGGAAGTGGTCTCGCGCTCAAGAGGCTTTCACCTTCCC 2700
Db 3339 GGCTGCCGTGGCAGGAAGCGGAAGTGGTCTCGCGCTCAAGAGGCTTTCACCTTCCC 3398
QY 2701 TGGAGGACCATCTGCCGTGATGAGCTGCAGAGCTGGGCTCGGTGAGAGTGAAGC 2760
Db 3399 TGGAGGACCATCTGCCGTGATGAGCTGCAGAGCTGGGCTCGGTGAGAGTGAAGC 3458
QY 2761 CTCCTGGAGACATCTCCCGGATCTTGAAGGCGAGCCCGAAGGTCTGGAGAGGAGG 2820

Db 3459 CTCCTGGAGCACATCTCCCGATCTTGAAGGCGAGGCCGGAAGGTCTGGAGAAGGAGG 3518
QY 2821 GCCCCCAGGAGAAAGCCAGGCTTGTCTTCCGCTCTCAGGTCTGAAGAGCTGGGA 2880
Db 3519 GCCCCCAGGAGAAAGCCAGGCTTGTCTTCCGCTCTCAGGTCTGAAGAGCTGGGA 3578
QY 2881 CCGAGCGCCGACATCTTCTAAGGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCA 2940
Db 3579 CCGAGCGCCGACATCTTCTAAGGGAGCTCTCAGATGAGACTGTGTCTTGGGCCAGTCA 3638
QY 2941 GACACTGGCTGCGAGGTGTGAGCCAGCCAGCTGCGCCAGGCGCACCTTGGAGCAAGACGG 3000
Db 3639 GACACTGGCTGCGAGGTGTGAGCCAGCCAGCTGCGCCAGGCGCACCTTGGAGCAAGACGG 3698
QY 3001 A 3001
Db 3699 A 3699

RESULT 14

US-10-921-168-3

; Sequence 3, Application US/10921168

; Publication No. US20050003446A1

; GENERAL INFORMATION:

; APPLICANT: WEI, Ming-Hui, et al

; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC

; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES

; FILE REFERENCE: CLO00927-CIP-DIV3

; CURRENT APPLICATION NUMBER: US/10/921,168

; PRIOR FILING DATE: 2004-08-19

; PRIOR APPLICATION NUMBER: 10/274,978

; PRIOR FILING DATE: 2002-10-22

; PRIOR APPLICATION NUMBER: 09/858,664

; PRIOR FILING DATE: 2001-05-17

; PRIOR APPLICATION NUMBER: 09/711,134

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 5207

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-921-168-3

Query Match 99.9%; Score 2997.8; DB 18; Length 5207;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CTTAGCGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTTGGCCGCGCTGAG 60
Db 699 CTTAGCGAGCAGAACTCGGGCCAGGCATACAGGAGCGAGACATCTTGGCCGCGCTGAG 758
QY 61 CCACCCGCTGGTCAAGGGCTCTGAGCCAGTTTGAACCCCGAAGACCTCATCTCAT 120
Db 759 CCACCCGCTGGTCAAGGGCTCTGAGCCAGTTTGAACCCCGAAGACCTCATCTCAT 818
QY 121 CTTGAGCTGTCTCATCGAGGAGCTGTGACCGCTGTACAGGAGGGGCTGGTGC 180
Db 819 CTTGAGCTGTCTCATCGAGGAGCTGTGACCGCTGTACAGGAGGGGCTGGTGC 878
QY 181 GGAGCCGAGGTCAAAGGTCTACATCCAGCAGCTGGTGGAGGGGCTGCATCTACCTGCACAG 240
Db 879 GGAGCCGAGGTCAAAGGTCTACATCCAGCAGCTGGTGGAGGGGCTGCATCTACCTGCACAG 938
QY 241 CCATGGCTTCTCCACCTGGACATAAGCCCTCTAAATCTGATGGTGCATCTTGGCCG 300
Db 939 CCATGGCTTCTCCACCTGGACATAAGCCCTCTAAATCTGATGGTGCATCTTGGCCG 998
QY 301 GGAAGACATTAATACTCGGACTTTGGCTTTGGCCAGAAATCATCCCCCAGCAGAGCTGCA 360
Db 999 GGAAGACATTAATACTCGGACTTTGGCTTTGGCCAGAAATCATCCCCCAGCAGAGCTGCA 1058

361 GTTCAGCAGTACGGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGCAGAACCCCTGT 420
1059 GTTCAGCAGTACGGCTCCCTGAGTTGCTCTCCCGAGATCATCCAGCAGAACCCCTGT 1118
421 GAGCGAAGCCTCCGACATTTGGGCCATATGGGTGTCTCTCTACCTCAGCCTGACCTGCTC 480
1119 GAGCGAAGCCTCCGACATTTGGGCCATATGGGTGTCTCTCTACCTCAGCCTGACCTGCTC 1178
481 ATCCCCATTTGCCGGCGAGAGTACCGTGCACACCTCTCTGAACTCTCTGGAGGGCGCGT 540
1179 ATCCCCATTTGCCGGCGAGAGTACCGTGCACACCTCTCTGAACTCTCTGGAGGGCGCGT 1238
541 GTATGAGCAGAGCCCATGGCTGCCACCTCAGCGAAGAGCCCAAGACTTCATCAAGGC 600
1239 GTATGAGCAGAGCCCATGGCTGCCACCTCAGCGAAGAGCCCAAGACTTCATCAAGGC 1298
601 TAGCCTCAGAGAGCCCATGAGCCCGGCTAGTGCAGCCAGTGCCTCTCCACCCCTG 660
1299 TAGCCTCAGAGAGCCCATGAGCCCGGCTAGTGCAGCCAGTGCCTCTCCACCCCTG 1358
661 GTTCCTGAATCCATGCTTGGGAGGAGGCCACCTTCATCAACCAAGCAGCTCAAGTT 720
1359 GTTCCTGAATCCATGCTTGGGAGGAGGCCACCTTCATCAACCAAGCAGCTCAAGTT 1418
721 CCTCTGGCCCGAAGTCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTCTGGTAT 780
1419 CCTCTGGCCCGAAGTCTGGCAGCGTTCCCTGATGAGCTACAAGTCCATCTCTGGTAT 1478
781 GCGCTCATCCTCAGCTGTGCGGGGCCCAACCGCAGACCCCTCCCTCGGCGTAGCCCG 840
1479 GCGCTCATCCTCAGCTGTGCGGGGCCCAACCGCAGACCCCTCCCTCGGCGTAGCCCG 1538
841 GCACCTTCAGAGGAGACTGTGTGCTCTCAGTTCTCTCTCTCTCTCTGACAAAGAGCT 900
1539 GCACCTTCAGAGGAGACTGTGTGCTCTCAGTTCTCTCTCTCTCTCTGACAAAGAGCT 1598
901 CGCCCTCATTTGCCCGGCTAAGTCACTGCCACCTCCCGGTGACACTCACCACCTGCT 960
1599 CGCCCTCATTTGCCCGGCTAAGTCACTGCCACCTCCCGGTGACACTCACCACCTGCT 1658
961 GCACCCCGGGGCTTCTCGGGCCCTCGCGCAGCTGCTGAGGAAGCCGAGGCGCAGTGA 1020
1659 GCACCCCGGGGCTTCTCGGGCCCTCGCGCAGCTGCTGAGGAAGCCGAGGCGCAGTGA 1718
1021 GCGCTCACAGAGCCCGAGCTCCGCTGCTATCTCCGAGGCTGCGGGCCACCGGCGC 1080
1719 GCGCTCACAGAGCCCGAGCTCCGCTGCTATCTCCGAGGCTGCGGGCCACCGGCGC 1778
1081 CCAGGGCTGTGCTGCCCGGACAGCTCATCCGAGCTGCTTACCAACAGGCGGGTGA 1140
1779 CCAGGGCTGTGCTGCCCGGACAGCTCATCCGAGCTGCTTACCAACAGGCGGGTGA 1838
1141 GAGCCTCAGCAGCGGGCCCTGGCCCGGGGAGCAGCGGCAACCGGCGCGGCGCA 1200
1839 GAGCCTCAGCAGCGGGCCCTGGCCCGGGGAGCAGCGGCAACCGGCGCGGCGCA 1898
1201 CTGCTGAAGGGCGGGTACATTGGGGGGGCGCTGCCAGGCTTGCAGGCACTGATGGA 1260
1899 CTTGCTGAAGGGCGGGTACATTGGGGGGGCGCTGCCAGGCTTGCAGGCACTGATGGA 1958
1261 GCACCGGCTGTGAGGAGGAGGCGCGCAGGAGGAGCGGCACCTCTCTGGCCAAAGC 1320
1959 GCACCGGCTGTGAGGAGGAGGCGCGCAGGAGGAGCGGCACCTCTCTGGCCAAAGC 2018
1321 CCCCTCATTTGAGACTGCGCTCGGCTGCTCTGAGCAGCCACCTTTGGCCCTGGCCA 1380
2019 CCCCTCATTTGAGACTGCGCTCGGCTGCTCTGAGCAGCCACCTTTGGCCCTGGCCA 2078
1381 CAGCCAATCCCTGAAATGATCTCCGAGCAACCCCGCCCTCTCTCGAGGCGCTGCGG 1440
2079 CAGCCAATCCCTGAAATGATCTCCGAGCAACCCCGCCCTCTCTCGAGGCGCTGCGG 2138
1441 TGAGGCACAGGACTGCTTCAGCCCCCTCGGGGGGGGGCCCTATATCAGGAGNATGSGGCA 1500

2139 TGAGGCACAGGACTGCTTCAGCCCCCTCGGGGGGGCCCTATCAGGAGNATGSGGCA 2198
1501 CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGTGTGGCCACCCAGGCACTGTCTCAGCCAGA 1560
2199 CCCTCAGGGCTCCAAAGCAGCTTCCATCCACTGTGTGGCCACCCAGGCACTGTCTCAGCCAGA 2258
1561 GAGGCCATCCCGGACAGCCCTTGGGGGCGACCCAGCCCTTCTTGCCACCCCAAGCAGGG 1620
2259 GAGGCCATCCCGGACAGCCCTTGGGGGCGACCCAGCCCTTCTTGCCACCCCAAGCAGGG 2318
1621 TTCTGCCCCCAGAGAGGCTGACGCCCCCAGCAGAGTTGCCCCCATGCCCCCTCTCTGGCTC 1680
2319 TTCTGCCCCCAGAGAGGCTGACGCCCCCAGCAGAGTTGCCCCCATGCCCCCTCTCTGGCTC 2378
1681 CTTCCTCCAGGATCTTTGCAAGAGGCCCTTATGTAACCTCAAGCCCTCTTTTGGGACA 1740
2379 CTTCCTCCAGGATCTTTGCAAGAGGCCCTTATGTAACCTCAAGCCCTCTTTTGGGACA 2438
1741 GCCCCAGGCAACCCCTGCCCCCTGCCAAGAGGCCCTTATGTAACCTCAAGAGTGGGGCC 1800
2439 GCCCCAGGCAACCCCTGCCCCCTGCCAAGAGGCCCTTATGTAACCTCAAGAGTGGGGCC 2498
1801 TGAGAGATCTCTCTTCTGGGAGGCCAAACCCGCGCCCTTGCAAGTTCCCCCAGGCTCAGC 1860
2499 TGAGAGATCTCTCTTCTGGGAGGCCAAACCCGCGCCCTTGCAAGTTCCCCCAGGCTCAGC 2558
1861 CTCCAGGCGAGCTTTTCCAAAGTGAAGTCTCTCAGGGTGGGCTCTCTCCAGGTGGGCA 1920
2559 CTCCAGGCGAGCTTTTCCAAAGTGAAGTCTCTCAGGGTGGGCTCTCTCCAGGTGGGCA 2618
1921 AGAGCTGGCCCCCTCTGAGTGGGAGGGCTGAGCCAGAGGCTGAGGATCTGTCCGA 1980
2619 AGAGCTGGCCCCCTCTGAGTGGGAGGGCTGAGCCAGAGGCTGAGGATCTGTCCGA 2678
1981 CTCCACACCACTTTCAGCGGCTCAGGAAACAGGCGACCAATGCGCAAGTTCTCCCTGGG 2040
2679 CTCCACACCACTTTCAGCGGCTCAGGAAACAGGCGACCAATGCGCAAGTTCTCCCTGGG 2738
2041 TGFTCGGGGGGTACCGCAGCGGTGTGCTATGCGCACTTTGCTTGTGTGAGATGC 2100
2739 TGFTCGGGGGGTACCGCAGCGGTGTGCTATGCGCACTTTGCTTGTGTGAGATGC 2798
2101 AGGGGCTGCTGGGGCAGGGGCCCATGTTGGCCAGGATAGCTGGGCTGTGTCCAGTC 2160
2799 AGGGGCTGCTGGGGCAGGGGCCCATGTTGGCCAGGATAGCTGGGCTGTGTCCAGTC 2858
2161 GGAGGAGGAGGAGCAGGAGGAGGCGAGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2220
2859 GGAGGAGGAGGAGCAGGAGGAGGCGAGGCTGAGTCCAGTCCGAGGAGCAGCAGGAGGC 2918
2221 CAGGGCTGAGAGGCCACTGCCCCCAGGTCAGTGCAGGCGCTGTGTCTGAGTCCGAGCGGC 2280
2919 CAGGGCTGAGAGGCCACTGCCCCCAGGTCAGTGCAGGCGCTGTGTCTGAGTCCGAGCGGC 2978
2281 TCCACACAGAGCTCTCCAGAGCCCAACCCATGAGGAGGACATCGGGCAGGCTCTCCCTGGT 2340
2979 TCCACACAGAGCTCTCCAGAGCCCAACCCATGAGGAGGACATCGGGCAGGCTCTCCCTGGT 3038
2341 GCAGATCCGGGACCTGTGCTGAGTGTGCGGAGGGCGCGACACAATATCCCTGGACATTTTC 2400
3039 GCAGATCCGGGACCTGTGCTGAGTGTGCGGAGGGCGCGACACAATATCCCTGGACATTTTC 3098
2401 CGAGGTGAGCCCGCCCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 2460
3099 CGAGGTGAGCCCGCCCTTACCTCAACCTCTCAGACCTGTACGATATCAAGTACCTCCCAT 3158
2461 CGAGTTTATGATCTTACGAAAAGTCCCCAAGTCCGCTCAGCAGAGCGGCCCTCCCCCAT 2520
3159 CGAGTTTATGATCTTACGAAAAGTCCCCAAGTCCGCTCAGCAGAGCGGCCCTCCCCCAT 3218
2521 GGCTGAGGAGGAGCTGCGGATTTCCCGGAGCCCACTGCGCCCTGGCCAGGTGAGTGGG 2580

Db 3219 GCGTGGAGGAGCTGGCCGAGTTCCGGAGGCCAGTGGCCCTGGCCAGGTGAACCTGGG 3278
QY 2581 CCCCCACGAGGCTGGAGATCACAGAGGAGTGGAGCGCTGTCGCGAGA 2640
Db 3279 CCCCCACGAGGCTGGAGATCACAGAGGAGTGGAGCGCTGTCGCGAGA 3338
QY 2641 GCGTGGCTGGGAGGAGGAGTGGTCTGGCGCTCACGAGCGCTTTCACCTTCCC 2700
Db 3339 GCGTGGCTGGGAGGAGGAGTGGTCTGGCGCTCACGAGCGCTTTCACCTTCCC 3398
QY 2701 TGGGAGGACCTGGCGCTGGAGGCTGGAGGCTGGGCTGGGCTGGAGAGTGAAGGC 2760
Db 3399 TGGGAGGACCTGGCGCTGGAGGCTGGAGGCTGGGCTGGGCTGGAGAGTGAAGGC 3458
QY 2761 CTCCTGGAGCACATCTCCCGGATCTCGAAGGCGAGGCGGGAAGTCTGGAGAGGAGGG 2820
Db 3459 CTCCTGGAGCACATCTCCCGGATCTCGAAGGCGAGGCGGGAAGTCTGGAGAGGAGGG 3518
QY 2821 GCCCCCCAGAGAGCCAGGCTTCTCTTCCCGCTCTCAGCTCTGAAGAGCTGGGA 2880
Db 3519 GCCCCCCAGAGAGCCAGGCTTCTCTTCCCGCTCTCAGCTCTGAAGAGCTGGGA 3578
QY 2881 CCGAGCGCCGACATCTCTAAGGAGCTCTCAGATGAGAGTGGTCTGGGCGGCTCAGT 2940
Db 3579 CCGAGCGCCGACATCTCTAAGGAGCTCTCAGATGAGAGTGGTCTGGGCGGCTCAGT 3638
QY 2941 GACACTGGCTCCAGGTGTGACCCAGCCAGCTGTCAGGCGGAGGAGGAGGAGGAGG 3000
Db 3639 GACACTGGCTCCAGGTGTGACCCAGCCAGCTGTCAGGCGGAGGAGGAGGAGGAGG 3698
QY 3001 A 3001
Db 3699 A 3699

RESULT 15
US-10-415-011-44
; Sequence 44, Application US/10415011
; Publication No. US20040053394A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: XU, Yuming
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: DING, Li
; APPLICANT: TANG, Y. Tom
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: NGUYEN, Dannel B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LU, Yan
; APPLICANT: YUE, Henry
; APPLICANT: BURFORD, Neil
; APPLICANT: BANDMAN, Olga
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: LAL, Preeti G.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: SWARNAKER Anita
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: KHAN, Farrah A.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0262 USN
; CURRENT APPLICATION NUMBER: US/10/415,011
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: PCT/US01/47728

Query Match 99.8%; Score 2996.2; DB 17; Length 5454;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2998; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CCTACGGAGCAGAACTCTGGGCGCCAGGCATACAGGGAGCGAGACATCTCTGGCGCGCTGAG 60
Db 963 CCTACGGAGCAGAACTCTGGGCGCCAGGCATACAGGGAGCGAGACATCTCTGGCGCGCTGAG 1022

QY 61 CACCCCGCTGGTCA CGGGGCTGCTGAG CCAAGTTCAGTTCAGTTCATCTCAT 120
Db 1023 CCACCGCTGGTCA CGGGGCTGCTGAG CCAAGTTCAGTTCAGTTCATCTCAT 1082

QY 121 CTTGAGAGCTGTCTCATCCGAGGAGCTGCTGAGCGGCTGTACAGAGGAGGCTGGTGC 180
Db 1083 CTTGAGAGCTGTCTCATCCGAGGAGCTGCTGAGCGGCTGTACAGAGGAGGCTGGTGC 1142

QY 181 GGAGGCCGAGGTCAAGGTCTACATCCAGCAGCTGGTGGAGGGGCTGCATCTCTGCACAG 240
Db 1143 GGAGGCCGAGGTCAAGGTCTACATCCAGCAGCTGGTGGAGGGGCTGCATCTCTGCACAG 1202

QY 241 CCATGGCGTCTCCACCTGGACATTAAGCCCTCTAACAATCTCTGATGGTGCATCTCTGCCG 300
Db 1203 CCATGGCGTCTCCACCTGGACATTAAGCCCTCTAACAATCTCTGATGGTGCATCTCTGCCG 1262

QY 301 GGAAGACATTAAATCTCGACTTTGGCTTTGGCCAGAACATCACCCAGAGAGCTGCA 360
Db 1263 GGAAGACATTAAATCTCGACTTTGGCTTTGGCCAGAACATCACCCAGAGAGCTGCA 1322

QY 361 GTTCAGCAGATACGGCTCCCTGAGTTCGTCTCCCGCAGATCATCCAGAGAGAACCTTGT 420
Db 1323 GTTCAGCAGATACGGCTCCCTGAGTTCGTCTCCCGCAGATCATCCAGAGAGAACCTTGT 1382

QY 421 GAGCGAAGCCTCCGACATTTGGGCCATGGGTGTCTCTCTTACCTCAGCTGACCTGCTC 480
Db 1383 GAGCGAAGCCTCCGACATTTGGGCCATGGGTGTCTCTCTTACCTCAGCTGACCTGCTC 1442

QY 481 ATCCCATTTTGC CGGCGAGAGTGACCGTGCACCCCTCTCTGAACCTCTCTGGAGGGGCGGT 540
Db 1443 ATCCCATTTTGC CGGCGAGAGTGACCGTGCACCCCTCTCTGAACCTCTCTGGAGGGGCGGT 1502

QY 541 GTCATGGAGAGCCCATGGCTGGCCACCTCAGGAGAGACCCCAAGACTTCATCAAGGC 600
Db 1503 GTCATGGAGAGCCCATGGCTGGCCACCTCAGGAGAGACCCCAAGACTTCATCAAGGC 1562

QY 601 TACGCTGAGAGAGCCCTCAGGCGCCGCTAGTGGCGCCGAGTGCCTCTCCCAACCCCTG 660
Db 1563 TACGCTGAGAGAGCCCTCAGGCGCCGCTAGTGGCGCCGAGTGCCTCTCCCAACCCCTG 1622

Accession	Sequence	Position
Db		3783
	GCCCCCAGGAAGACGCGCTTGTTCTCTCCGCTCTCAGGCTGGAAGAGCTGGGA	3842
Qy		2881
	CCGAGCCCGACATTCCTAAGGGAGCTCTCAGATGAGACTGTGTCTCTGGCCAGTCAGT	2940
Db		3843
	CCGAGCCCGACATTCCTAAGGGAGCTCTCAGATGAGACTGTGTCTCTGGCCAGTCAGT	3902
Qy		2941
	GACACTGGCCTGCCAGCTGTTCAGCCAGCCAGCTGCCAGGCCACTCTGGAGCAAGACGG	3000
Db		3903
	GACACTGGCCTGCCAGCTGTTCAGCCAGCCAGCTGCCAGGCCACTCTGGAGCAAGACGG	3962
Qy		3001 A 3001
		3963 A 3963
Db		

Search completed: March 21, 2005, 16:08:22
Job time : 1893.9 secs

301 AGAGATGGGACAGATACAGCTGAGGACGAGCGGACTAAATGTGAGCTGCAGATTCTG 360
|||||
10800 AGAGATGGGACAGATACAGCTGAGGACGAGCGGACTAAATGTGAGCTGCAGATTCTG 10859
|||||
361 GGCCTGGCCATGGCAGACACTGGGAGTACTCGTGTGCGGACGAGAGACCTCG 420
|||||
10860 GGCCTGGCCATGGCAGACACTGGGAGTACTCGTGTGCGGACGAGAGACCTCG 10919
|||||
421 GCTATGCTCAGCTCAGGCTCTACCCATCAAGTTCACAGAGGGTCTCAGGAACGAAG 480
|||||
10920 GCTATGCTCAGCTCAGGCTCTACCCATCAAGTTCACAGAGGGTCTCAGGAACGAAG 10979
|||||
481 GCCACAGAGGGGCAACAGCGCTGCTGGGTGTGAGCTGAGCAAGATGSCCCCGTGGAG 540
|||||
10980 GCCACAGAGGGGCAACAGCGCTGCTGGGTGTGAGCTGAGCAAGATGSCCCCGTGGAG 11039
|||||
541 TGGTGGAAAGGGGCATGAGACCTCTCAGAGATGGAGACAGACAGCTCAGGACGAGCGG 600
|||||
11040 TGGTGGAAAGGGGCATGAGACCTCTCAGAGATGGAGACAGACAGCTCAGGACGAGCGG 11099
|||||
601 GCCAGGTGTGAGCTCAGATCGCGGCTCTGTCGAGAGAGAGCGCTGGGGAGTACTCTGTGC 660
|||||
11100 GCCAGGTGTGAGCTCAGATCGCGGCTCTGTCGAGAGAGAGCGCTGGGGAGTACTCTGTGC 11159
|||||
661 ATGTGCGGGAAGGAGGAGGACCTCAGCCATGCTCACCGTCAGGGCCATCGCTTCCAAGTTC 720
|||||
11160 ATGTGCGGGAAGGAGGAGGACCTCAGCCATGCTCACCGTCAGGGCCATCGCTTCCAAGTTC 11219
|||||
721 ATAGAGGGTCTGAGGAATGAAGAGGCCACAGAGGGGACAGCGCCACCGCTGTGTGTGAG 780
|||||
11220 ATAGAGGGTCTGAGGAATGAAGAGGCCACAGAGGGGACAGCGCCACCGCTGTGTGTGAG 11279
|||||
781 CTGAGCAAGGGCGGACCCGGTGTGAGAGAGAGAGGGGACATGAGACCTCAGAGATGGGAC 840
|||||
11280 CTGAGCAAGGGCGGACCCGGTGTGAGAGAGAGAGGGGACATGAGACCTCAGAGATGGGAC 11339
|||||
841 AGACACACCTTGAGGACGAGCGGTCCAGGTGTGAGCTGAGATCCGTGGGCTGCTGTG 900
|||||
11340 AGACACACCTTGAGGACGAGCGGTCCAGGTGTGAGCTGAGATCCGTGGGCTGCTGTG 11399
|||||
901 GTGATGCGGGGAGTACTCGTGTGTGTGTCGGGACGAGAGACCTCAGCACACTCACT 960
|||||
11400 GTGATGCGGGGAGTACTCGTGTGTGTGTCGGGACGAGAGACCTCAGCACACTCACT 11459
|||||
961 GTCAGGGCCCTGCTGCCAGATTCATAGAAGATGTGAAAA 1001
|||||
11460 GTCAGGGCCCTGCTGCCAGATTCATAGAAGATGTGAAAA 11500
|||||

RESULT 2
AB046776
LOCUS AB046776 4789 bp mRNA linear PRI 22-FEB-2001
DEFINITION Homo sapiens mRNA for KIAA1556 protein, partial cds.
ACCESSION AB046776
VERSION AB046776.1 GI:10047176
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)
REFERENCE
AUTHORS Negase,T., Kikuno,R., Nakayama,M., Hirose,M. and Ohara,O.
TITLE Prediction of the coding sequences of unidentified human genes.
XVII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro
JOURNAL DNA Res. 7 (4), 273-281 (2000)
MEDLINE 20450683
PUBMED 10997877
REFERENCE
AUTHORS Ohara,O., Negase,T. and Kikuno,R.
TITLE Direct Submission
JOURNAL Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba

292-0812, Japan (E-mail:cdnainfo@kazusa.or.jp,
URL:http://www.kazusa.or.jp/huge, Tel:81-438-52-3913,
Fax:81-438-52-3914)
FEATURES
source
Location/Qualifiers
1. 4789
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fhl18619"
/note="vector:pBluescriptII SK plus"
1. 4789
/gene="KIAA1556"
1. 4789
/gene="KIAA1556"
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1556 protein"
/protein_id="BAB13382.1"
/db_xref="GI:10047177"
/translation="RAVRVHDLHVGITKRLKTNLELEGESCECVLSHESASDPM
WTGKTVGSSRFQATRGKRIYLVREAPSDAGEVVFVSRGLTSKASLIVRERPA
AIIKPLEDQWAPGEDVELRCELSRAGTPVHMLKDKRAIKRSQIDYVVECTMAVLVI
RGASLKDAGEYTCVEASKSTASLHVEEKANCTFEELTNLQVEKGTAVFTCKTEHPA
ATVWRKGLLELRASGHQSQEGLRLTLTSALEKADSDTYTCDIGQAQRAQLVQ
GRRVH11EDLEDVDVQSSSATPRCISIPANYEPVHFLDKTPIHANELNEDIDAPGG
YHVLTLRQALAKDSGTIYFAGDQDASALRLVTEKPSFSELTDAITTEGEDILTVC
ETSDCDIPVMTKDGKTLRSARCOLSHEHRAQLITGATLQDSGRYKCBAGACSS
SIVRHARPVRFQALDXLDLEVEGAATLRCVLSVAAPVKMGVNNVLRPGDKYSLR
QEGMALVLRNLRFQDSGRYSCSGQDTTATLVTALPAQFICKLRNKEATGATA
TLRCELSKAAPEVWRKGETLRDGDYCLRDQGANCELOIRGLAMVDAEYSCVCGEE
RTSASLTIRWPAHFIQRLRHOESIEGATATLRCELSKAAPEVWRKGRSLRDGRHS
LRQGANVCELOICGLAVADAGEYSCVCGEERTSATLTVKALPAKFTGLRNEAVEGA
TAMUWCELSKAAPEVWRKGPENLRDGRYILRQEGTRCELOICGLAMADAGEYLCVCG
QERTSATLTTRALPARFIEDVKNQAREGATVLOCELSNAAPEVWRKGSBTLRDGR
YSLRQDGTCELOIRGLAMADGETYSCVCGEERTSATLTTRALPIKFTEGRLNEATE
GATLAVCELSKAAPEVWRKHETLRDGRHSLRQDGRARCELOIRGLVADAGEYLCM
CGKERTSAMLTVRAMPSKFTGLRNEATEGDTATLMCELSKAAPEVWRKHETLRDGR
RHSLRDGRSCELOIRGLAVADAGEYSCVCGEERTSATLTTRALPARFIEDVKNQEA
REGATVLOCELSKAAPEVWRKGETLRDGRYILRQDGTTRCELOIHGLSVADTGEYS
CVCQERTSATLTTRALPARPTDQKTEASEGATATLQCELSKAAPEVWRKGPETLR
DGRYSLRQDGTTRCELOIHGLSVADAGEYSCVCGEERTSATLTTRALPARFTGLRNE
EAMGATATLQCELSKAAPEVWRKGEALRDGRYSLRQDGTTRCELOIRGLAMADGV
YSCVCGEERTSATLTTRALPARFIEDVKNQAREGATVLOCELSKAAPEVWRKGPNT
LKDGRYSLRQDGTTRCELOIRGLVADAGEYSCVCGEERTSATLTTRALPARFIEDVR
NEATEGATVLOCELSKAAPEVWRKGETLRDGRYSLRQDGTTRCELOIRGLAVEDT
GEYLCVCGEERTSATLTTRALPARFIEDVKNQAREGATATLHCELSKAAPEVWRKGP
ETLRDGRHSLRQENLRNLPNGGGGSELGSCSPAWTAKLDFIS"

ORIGIN

Query Match 99.7%; Score 997.8; DB 9; Length 4789;
Best Local Similarity 99.8%; Pred. No. 1.2e-20;
Matches 999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCCCTGTGGAGTGGAGNAGGGCCCGAGAACCTCAGAGATGGGACAGATACATCTCTG 60
DB 2256 GCCCTGTGGAGTGGAGNAGGGCCCGAGAACCTCAGAGATGGGACAGATACATCTCTG 2315
QY 61 AGGCAGAGGGGACAGGTGTGAGCTGTGCGCTTGTGCGCATGGCCGACGCCCGG 120
DB 2316 AGGCAGAGGGGACAGGTGTGAGCTGTGCGCTTGTGCGCATGGCCGACGCCCGG 2375
QY 121 GAGTACTGTGTGTGGGGCAGAGAGGACCTCAGCCACGCTCACCATCAGGGCTCTG 180
DB 2376 GAGTACTGTGTGTGGGGCAGAGAGGACCTCAGCCACGCTCACCATCAGGGCTCTG 2435
QY 181 CCTGCCAGTTTCATAGAAGATGTGAAAAACAGAGAGCCAGAGAGGGGCGCACCGCTGTG 240
DB 2436 CCTGCCAGTTTCATAGAAGATGTGAAAAACAGAGAGCCAGAGAGGGGCGCACCGCTGTG 2495
QY 241 CTGCAGTGTGAGCTGAACAGTGCAGCCCTCTGTGGAGTGGAGAAAGGGGTCTTGAGACCTTC 300
DB 2496 CTGCAGTGTGAGCTGAACAGTGCAGCCCTCTGTGGAGTGGAGAAAGGGGTCTTGAGACCTTC 2555

QY 301 AGAGATGGGACAGATACAGCTTGAGCGAGGACCGGACTAAATGTGAGCTGCAGATTCTGT 360
Db AGAGATGGGACAGATACAGCTTGAGCGAGGACCGGACTAAATGTGAGCTGCAGATTCTGT 2615
QY 361 GGCTGGCCATGGCAGACACTGGGAGTACTGTGGTGTGGGCGAGGAGGACCTGC 420
Db GGCTGGCCATGGCAGACACTGGGAGTACTGTGGTGTGGGCGAGGAGGACCTGC 2675
QY 421 GCTATGCTCACTCAGGGCTCTACCCATCAAGATTACAGAGGGTCTGAGGAACCAAGAG 480
Db GCTATGCTCACTCAGGGCTCTACCCATCAAGATTACAGAGGGTCTGAGGAACCAAGAG 2735
QY 481 GCCAAGAGGGGCAACACGGTGTGCGGTGTGAGCTGAGCAAGATGSCCCCGTGGAG 540
Db GCCAAGAGGGGCAACACGGTGTGCGGTGTGAGCTGAGCAAGATGSCCCCGTGGAG 2795
QY 541 TGTGGAAGGGGCAACACGGTGTGCGGTGTGAGCTGAGCAAGATGSCCCCGTGGAG 600
Db TGTGGAAGGGGCAACACGGTGTGCGGTGTGAGCTGAGCAAGATGSCCCCGTGGAG 2855
QY 601 GCCAGGTGTGAGCTGCAGATCCGCGGCTCTGCGGAGAGGACCGCTGGGAGTACCTGTGC 660
Db GCCAGGTGTGAGCTGCAGATCCGCGGCTCTGCGGAGAGGACCGCTGGGAGTACCTGTGC 2915
QY 661 ATGTGGGGAAGCAGAGGACCTCAGCATGCTCAACGCTCAGGCGCATGCTTCCAAAGTTC 720
Db ATGTGGGGAAGCAGAGGACCTCAGCATGCTCAACGCTCAGGCGCATGCTTCCAAAGTTC 2975
QY 721 ATAGAGGGTCTCAGAAATGAGAGGCCACAGAGGGGACCGCTGGGAGTACCTGTGTGAG 780
Db ATAGAGGGTCTCAGAAATGAGAGGCCACAGAGGGGACCGCTGGGAGTACCTGTGTGAG 3035
QY 781 CTGACAAAGGGGCAACCGGTGAGTGGAGGAGGCGCATGAGACCTCAGAGATGGGGAC 840
Db CTGACAAAGGGGCAACCGGTGAGTGGAGGAGGCGCATGAGACCTCAGAGATGGGGAC 3095
QY 841 AGACACGCTTGAGCAGACGGTTCAGGTGTGAGCTGAGATCCGTGGCTGGCTGTG 900
Db AGACACGCTTGAGCAGACGGTTCAGGTGTGAGCTGAGATCCGTGGCTGGCTGTG 3155
QY 901 GTGGATGCGGGGAGTACTCGTGTGTGTGCGGGCAGGAGGACCTCAGGCACACTCACT 960
Db GTGGATGCGGGGAGTACTCGTGTGTGTGCGGGCAGGAGGACCTCAGGCACACTCACT 3215
QY 961 GTCAGGCGCTGCTGCCAGATTATAGAAAGATGTGAAAA 1001
Db GTCAGGCGCTGCTGCCAGATTATAGAAAGATGTGAAAA 3256

RESULT 3
LOCUS CQ730656 18524 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 16590 from Patent W002068579.
ACCESSION CQ730656
VERSION CQ730656.1 GI:42305092
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kits, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 16590 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source
1..18524
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Query Match 71.3%; Score 714; DB 6; Length 18524;
Besc Local Similarity 83.5%; Pred. No. 3.le-142;
Matches 810; Conservative 0; Mismatches 160; Indels 0; Gaps 0;
QY 1 GCCCTGTGGAGTGGAGAGGGGCCCGAGAACTCTCAGAGATGGGACAGATACATCCTG 60
Db GCCCTGTGGAGTGGAGAGGGGCCCGAGAACTCTCAGAGATGGGACAGATACATCCTG 11142
QY 61 AGGCGAGAGGGGACCAAGGTGTGAGCTCAGATCTGTGCGCTTGGCCATGGCGGACGCGGG 120
Db AGGCGAGAGGGGACCAAGGTGTGAGCTCAGATCTGTGCGCTTGGCCATGGCGGACGCGGG 11202
QY 121 GAGTACTGTGTGTGCGGGCAGGAGGACCTCAGCAGCTCACCATCAGGGTCTGT 180
Db GAGTACTGTGTGTGCGGGCAGGAGGACCTCAGCAGCTCACCATCAGGGTCTGT 11262
QY 181 CCTGCCAGGTTCATAGAAAGATGTGAAAAACACAGAGGCCACAGAGAGGGGCCACGGCTGTG 240
Db CCTGCCAGGTTCATAGAAAGATGTGAAAAACACAGAGGCCACAGAGAGGGGCCACGGCTGTG 11322
QY 241 CTGCAGTGTGAGCTGAACAGTGCAGCCCTGTGTGAGTGGAGAAAGGGGTCTGAGACCCCTC 300
Db CTGCAGTGTGAGCTGAACAGTGCAGCCCTGTGTGAGTGGAGAAAGGGGTCTGAGACCCCTC 11382
QY 301 AGAGATGGGACAGATACAGCTTGAGGAGGACGGGAGCTGAGCTGCAGATTCTGT 11501
Db AGAGATGGGACAGATACAGCTTGAGGAGGACGGGAGCTGAGCTGCAGATTCTGT 11442
QY 361 GGCTTGCCCATGTCAGACACTGGGGAGTACTCGTGTGTGTGCGGGCAGGAGAGACCTCG 420
Db GGCTTGCCCATGTCAGACACTGGGGAGTACTCGTGTGTGTGCGGGCAGGAGAGACCTCG 11502
QY 421 GCTATGCTCACTCAGGGCTCTACCCATCAAGTTTACAGAGGGTCTGAGGAACGAGAG 480
Db GCTATGCTCACTCAGGGCTCTACCCATCAAGTTTACAGAGGGTCTGAGGAACGAGAG 11562
QY 481 GCCACAGAAAGGGGCAACAGCCGTGCTGCGGTGTGAGCTGAGCAAGATGCCCCCGTGGAG 540
Db GCCACAGAAAGGGGCAACAGCCGTGCTGCGGTGTGAGCTGAGCAAGATGCCCCCGTGGAG 11622
QY 541 TGTGGAAGGGGCACTGAGACCTCTCAGATGGAGACAGACAGCTGAGGCGAGGACGGG 600
Db TGTGGAAGGGGCACTGAGACCTCTCAGATGGAGACAGACAGCTGAGGCGAGGACGGG 11682
QY 601 GCCAGGTGTGAGCTGCAGATCCGCGGCTCTGCGGAGAGGACGCTGGGAGTACCTGTGC 660
Db GCCAGGTGTGAGCTGCAGATCCGCGGCTCTGCGGAGAGGACGCTGGGAGTACCTGTGC 11742
QY 661 ATGTGCGGAAAGGAGAGGACCTCAGCCATGCTCACCGTCAGGGGCCATGCTTCCAAAGTTC 720
Db ATGTGCGGAAAGGAGAGGACCTCAGCCATGCTCACCGTCAGGGGCCATGCTTCCAAAGTTC 11802
QY 721 ATAGAGGGTCTGAGGAATGAAGGGCCACAGAGGGGACACGGCCACGCTGTGTGTGAG 780
Db ATAGAGGGTCTGAGGAATGAAGGGCCACAGAGGGGACACGGCCACGCTGTGTGTGAG 11862
QY 781 CTGAGCAAGCGCGGCAACCGGTGAGTGGAGAGGGGACATGAGACCTCAGAGATGGGGAC 840
Db CTGAGCAAGCGCGGCAACCGGTGAGTGGAGAGGGGACATGAGACCTCAGAGATGGGGAC 11922
QY 841 AGACACGCTTGAGGAGGACCGGCTCAGGTTGAGTGCAGATCCGTGGCTGTGTG 900
Db AGACACGCTTGAGGAGGACCGGCTCAGGTTGAGTGCAGATCCGTGGCTGTGTG 11982
QY 901 GTGGATGCGGGGAGTACTCGTGTGTGTGCGGGCAGGAGAGGACCTCAGCCACACTCACT 960
Db GTGGATGCGGGGAGTACTCGTGTGTGTGCGGGCAGGAGAGGACCTCAGCCACTCACT 12042
QY 961 GTCAGGCGCC 970
Db GTCAGGCGCC 12111


```
RESULT 4
AX069754/c
LOCUS
DEFINITION Sequence 226 from Patent WO0102568.
ACCESSION AX069754
VERSION AX069754.1 GI:12579539
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Williams, L.T., Escobedo, J., Innis, M.A., Garcia, P.D., Klinger, J.,
Kasam, A., Reinhard, C., Randazzo, F., Kennedy, G.C., Pot, D.,
Lansom, G., Drmanac, R., Crkenjakov, R., Drmanac, S., Dickson, M.,
Labat, I., Lezhkowitz, D., Kita, D., Garcia, V. and Strache-Crain, B.
HUMAN GENES AND GENE EXPRESSION PRODUCTS
Patent: WO 0102568-A 226 11-JAN-2001.
JOURNAL CHIRON CORPORATION (US); HYSEQ, INC. (US)
FEATURES
source
1. 352
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
misc_feature
1..352
/note="n = A, T, C or G"
ORIGIN
Query Match 34.3%; Score 343.4; DB 6; Length 352;
Best Local Similarity 99.7%; Pred. No. 6.9e-63;
Matches 344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 631 GTGGCAGAGGACGCTGGGGAGTACCTGTGCATGTGCGGGAAGGAGGACCTCAGCCATG 690
Dy
Dy 351 GTGGCAGAGGACGCTGGGGAGTACCTGTGCATGTGCGGGAAGGAGGACCTCAGCCATG 292
Oy 691 CTCACCGTCAGGCGCATGCCCTCCAGTTTCATAGAGGTCTGAGGAATGAGAGGCCACA 750
Dy 291 CTCACCGTCAGGCGCATGCCCTCCAGTTTCATAGAGGTCTGAGGAATGAGAGGCCACA 232
Oy 751 GAAGGGGACAGCGCCAGCTGTGGTGTGAGCTGAGCAAGCGGCGCACCGTGGAGTGGAGG 810
Dy 231 GAAGGGGACAGCGCCAGCTGTGGTGTGAGCTGAGCAAGCGGCGCACCGTGGAGTGGAGG 172
Oy 811 AAGGGGATGAGACCTTCAGAGATGGGACAGACACAGCCTCAGGACGAGCGGTCCAGG 870
Dy 171 AAGGGGATGAGACCTTCAGAGATGGGACAGACACAGCCTCAGGACGAGGTCCAGG 112
Oy 871 TGTGAGCTGAGATCCGTTGGCTGGCTGTGAGTGGATGCCGGGGAGATCTCGTGTGTGTGC 930
Dy 111 TGTGAGCTGAGATCCGTTGGCTGGCTGTGAGTGGATGCCGGGGAGTCTCGTGTGTGTGC 52
Oy 931 GGGCAGGAGAGGACCTCAGGCACACTCACTCTCAGGGCCCTGCCT 975
Dy 51 GGGCAGGAGAGGACCTCAGGCACACTCACTCTCAGGGCCCTGCCT 7
RESULT 5
HSA314905
LOCUS
DEFINITION Homo sapiens partial OBSCN gene for obscurin, exons 31-56.
ACCESSION AJ314905
VERSION AJ314905.1 GI:21104337
KEYWORDS OBSCN gene; obscurin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1. Young, P., Ehler, B. and Gautel, M.
Obscurin, a giant sarcomeric Rho-GEF protein involved in sarcomere
assembly
```

```
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 34667)
AUTHORS Gautel, M.S.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2001) Gautel M.S., Physikalische Biochemie,
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse
11, Dortmund, 44227, GERMANY
FEATURES
source
1. 34667
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
247..34433
/gene="OBSCN"
join(247..519,1332..1598,1817..2082,4388..4664,5087..5353,
6008..6274,6375..6638,6939..7202,10212..10475,10837..11100,
11667..11930,12466..12729,13094..13357,16768..17031,
17573..17836,18196..18459,21989..22252,22617..22880,
23476..23700,24686..24952,25215..25481,25573..25846,
26423..26697,27413..27609,27798..27873,34161..34433)
/gene="OBSCN"
join(<247..519,1332..1598,1817..2082,4388..4664,5087..5353,
6008..6274,6375..6638,6939..7202,10212..10475,10837..11100,
11667..11930,12466..12729,13094..13357,16768..17031,
17573..17836,18196..18459,21989..22252,22617..22880,
23476..23700,24686..24952,25215..25481,25573..25846,
26423..26697,27413..27609,27798..27873,34161..34433)
/gene="OBSCN"
/codon_start=3
/product="obscurin"
/protein_id="CAC85752.1"
/db_xref="GI:21104338"
/db_xref="UniProt/Tx:EMBL:OBHNM3"
/t_xref="LHGVI:KRLKTMVLEGECSFECVLSHESASDPAMTVGKTV
GSSRFQATQGRKYILVRENAPSDAGEVVSURGLTSKASLIVRERPAIILPLED
QWAPGDEVELRCELSRAGTPVHKWKORKAIRKSKQKDVCEGTWMLVIRGASLKD
GEYTCVEASKSTASLHVKNKCTEELNLOVEKGTAVFTCKTBEHPATVTKRKG
LLELRASGKHPQSQEGLTLRLTISALEKADSDTYTCDIGQASQALLVQARRVHII
LELDVDVQSGSATFCRISPNAYEPVHFWLTKPLHANELNBIADQPGVHVLTLR
QVAKDGTITFEAGDORASALRVTEKPSVRELTDTATTEGEDTLVCTESTCDI
PLWTKDGTILRGARQLSHGHRAQLITGATLQDSGRYKCSAGGACSSIVRHA
RPVRFQALDKLEVEGATLRCVLSVAAPVKWCYGNVLRFDKYSLRQEGNMLE
LVNLRPQDSGRYSCFGDQTSATLTVLPAQFICKLNKATEGATLRCELS
KAAPVWRKGSSETLRDQDCLRGDGMCELQIRGLVADAAEYSCVGEERTSALT
IRPMPAFHIGRLRHOESIEGATLRCELSKAAPVWRKGSRLDRGHSRLRDGAV
CELQICGLAVADAGEYSCVGEERTSATLTVKALPAKFTGLRNEEA VEGATAMLCE
LSKAPVWRKGPENLRDGRYILRQSGTRCELQICGLAVADAGEYLCVCGQERTSAT
LTVRALPARFIEDVKNQAREGATVLCLELNSAAPVWRKGSSETLRDQDGRYSLRQDG
TKCELQIRGLVADAGEYSCVCGQERTSATLTVRALPFIKFTGLRNEEAETGATVLR
CELSKMAAPVWRKGHETLRDGRHSRLRDQDGCARCELQIRGLVADAGEYLCMCKERTS
AMLTVRAMPKSFIEGLRNEEATGDTATLWCELSKAAAPVWRKGHETLRDGRHSRLRQ
DGRCCELQIRGLVADAGEYSCVCGQERTSATLTVRALPARFIEDVKNQAREGATV
LQCELSKAAAPVWRKGSSETLRDGRYSLRQDGTCELQIHGLSVADTGEYSCVCGQER
TSATLTVRALPARFIDQKTRKEASEGATLQCELSKAPVWRKGPENLRDGRYSLRQDG
TKQDGTCELQIHGLSVADAGEYSCMCGQERTSATLTVRALPARFTEGLRNEEAETGAT
ATQCELSKAAAPVWRKGSSETLRDGRYSLRQDGCARCELQIRGLVADAGEYSCVCGQ
ERTSATLTVRALPARFIEDVKNQAREGATVLCLELNSAAPVWRKGSSETLRDQDGRYSLRQDG
SLKQDGTSCCELQIRGLVADAGEYSCICEOERTSATLTVRALPARFIEDVKNHATEG
ATVAVLQCELSKAAAPVWRKGSSETLRDGRYSLRQDGTCELQIRGLVADTGEYLCVC
QERTSATLTVRALPARFIDNMTQAREGATLHCELSKAPVWRKGPENLRDGRYSLRQDG
RHSRLRDGTCELQIRGLSVADAGEYSCVCGQERTSATLTVRTEGATGATMLQCELSK
KAPVWRKGPENLRDGRYSLRQDGTCELQIHGLSVADTGEYSCVCGQERTSATLTV
KAPQVREPLQSLQAEAGTATLQCELSBPTATVWMSKGLQVQAGREPRQCGCT
AELVLQDQREDTGEYTCGQSATSATLTVTAAPVRLRQHQVEDGEGTAHLCCCE
LSRAGSVWRKGSLSQLPCKAKYQWODGAAELLVRGVEQEDAGDCTDGTGHTQSM
SLSVRPRPKTRLRKLEQEDGIALCCQLSDAESGAVQVWLKGLVADAGEYLCVCGQER
RSQGTARELLIHQLEAKDTGEYACVTCGQKTAASLRVTGEPEVTIRGLVADTGEY
DVFECSVSRAGATGVQWCLQPLQNEVTEAVRDRGIRHLKRLKGTTPDADGTAF
HLGNHASSAQLTVRAPEVTILEPLQDQSEGODASFCQRLSRAGSGAGKRWALGVPL
QANEMNDITVQEGTLHLTKVLTLEDAGTVSFHVGTCSSQAQKTKATKNTVRLGLE
VEALEGGEALFECQLSQPEVNAHTWLLDDEPVHVSNAEVFFENGLRHLILLNLRP
```


exon	QDSCRVTFLAGDMVTSFAFLTVR 247. .519 /gene="OBSCN" /number=31	exon	13094. .13357 /gene="OBSCN" /number=43
intron	520. .1331 /gene="OBSCN" /number=31	intron	13358. .16767 /gene="OBSCN" /number=43
exon	1332. .1598 /gene="OBSCN" /number=32	exon	16768. .17031 /gene="OBSCN" /number=44
intron	1599. .1816 /gene="OBSCN" /number=32	intron	17032. .17572 /gene="OBSCN" /number=44
exon	1817. .2082 /gene="OBSCN" /number=33	exon	17573. .17836 /gene="OBSCN" /number=45
intron	2083. .4387 /gene="OBSCN" /number=33	intron	17837. .18195 /gene="OBSCN" /number=45
exon	4388. .4664 /gene="OBSCN" /number=34	exon	18196. .18459 /gene="OBSCN" /number=46
intron	4665. .5086 /gene="OBSCN" /number=34	intron	18460. .21988 /gene="OBSCN" /number=46
exon	5087. .5353 /gene="OBSCN" /number=35	exon	21989. .22252 /gene="OBSCN" /number=47
intron	5354. .6007 /gene="OBSCN" /number=35	intron	22253. .22616 /gene="OBSCN" /number=47
exon	6008. .6274 /gene="OBSCN" /number=36	exon	22617. .22880 /gene="OBSCN" /number=48
intron	6275. .6374 /gene="OBSCN" /number=36	intron	22881. .23475 /gene="OBSCN" /number=48
exon	6375. .6638 /gene="OBSCN" /number=37	exon	23476. .23700 /gene="OBSCN" /number=49
intron	6639. .6938 /gene="OBSCN" /number=37	intron	23701. .24685 /gene="OBSCN"
exon	6939. .7202 /gene="OBSCN" /number=38	<div>Query Match 33.5%; Score 335.2; DB 9; Length 34667; Best Local Similarity 66.1%; Pred. No. 1.7e-61; Matches 549; Conservative 0; Mismatches 243; Indels 38; Gaps 3;</div>	
intron	7203. .10211 /gene="OBSCN" /number=38	QY 175	GCTCTGCCTGCCAGGTTTCATAGAAGATGTGAAAAACACAGAGGCCAGAGAGGGGCCACG 234
exon	10212. .10475 /gene="OBSCN" /number=39	Db 6374	GCCTGCTGCCCGGAGTTTCATCGGAAACTGAGAAAACAGGAGGCCACAGAGGGGCCACG 6433
intron	10476. .10836 /gene="OBSCN" /number=39	QY 235	GCTGTGCTGCAGTGTGAGCTGAACAGTGCAGGCCCTGTGGAGTGGAGAAAGGGGTCTGAG 294
exon	10837. .11100 /gene="OBSCN" /number=40	Db 6434	GCCACGCTCGGTGTGAGTGTGAGCTGAGCAAGGAGGCCCTGTGGAGTGGAGAAAGGGGTCCGAG 6493
intron	11101. .11666 /gene="OBSCN" /number=40	QY 295	ACCTCAGAGATGGGGACAGATACAGCCTGAGGCAGGACGGGACTAAATGTGAGCTGCAG 354
exon	11667. .11930 /gene="OBSCN" /number=41	Db 6494	ACCCTCAGAGATGGGGACAGATACAGCCTGAGGCAGGACGGGACTAAATGTGAGCTGCAG 354
intron	11931. .12465 /gene="OBSCN" /number=41	QY 355	ATTCGTGGCTGGCCATGSCACACACTGGGGAGTACTCGTGTGTGCGGGCAGAGAGG 414
exon	12466. .12729 /gene="OBSCN" /number=42	Db 6554	ATCCGTGGCTGGCCATGSGTGGATGCCGCGGAGTACTCGTGTGTGTGTGGAGAGAGG 6613
intron	12730. .13093 /gene="OBSCN" /number=42	QY 415	ACCTGGCTATGCTCACCGTCAAGGCTCTACCCATCAAGTTCCACAGAGGGTCTGAGGAAC 474
		Db 6614	ACCTCAGCCTCACTCACCATCAGGCGTAAGACCGTGTATCCAGAGCCGTGTCCGTGTCC 6673
		QY 475	GA-----AGAGGCCACAGAAAGGGGCAACAGCC 501
		Db 6674	AAATTTTACCTCTGCTGTGAGTCCGTACCCATTCCTAGTCTACCTGACTGTGCCATGCC 6733
		QY 502	GTGCTGCGGTGTGAGCTGAGCAAGATGCCGCCGCGGAGTGTGTGAAGGGGATGAGACC 561
		Db 6734	TTCCACAGCACATCAGGGAATCCCTTCAGACTTTCTGGGCTGTGGAAGGAACCGAGGAAC 6793


```
OY      918 CTCGTGTGTGTCGGGACGAGGACCTCAGCCACATCACTGTCAGGG 967
Db      76166 CTCCTGTGTGTGCGGAGGAGACCTCTGCCACTCTCCACGTGAAGG 76215

RESULT 7
AC026657
LOCUS   AC026657      164766 bp    DNA    linear    HTG 01-SEP-2000
DEFINITION Homo sapiens chromosome 1 clone RP11-245P10, WORKING DRAFT
SEQUENCE, 31 unordered pieces.
ACCESSION AC026657
VERSION   AC026657.4 GI:9958202
KEYWORDS  HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 164766)
REFERENCE 1
AUTHORS   Waterston,R.H..
TITLE     The sequence of Homo sapiens clone
JOURNAL   Unpublished
REFERENCE 2 (bases 1 to 164766)
AUTHORS   Waterston,R.H..
TITLE     Direct Submission
JOURNAL   Submitted (22-MAR-2000) Genome Sequencing Center, Washington
          University School of Medicine, 4444 Forest Park Parkway, St. Louis,
          MO 63108, USA
COMMENT   On Sep 1, 2000 this sequence version replaced gi:7637349.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0245P10
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Assembly: Dye-terminator Big Dye; 0% of reads
Consensus program: Phrap; version 0.990319
Consensus quality: 139884 bases at least Q40
Consensus quality: 147686 bases at least Q30
Consensus quality: 151469 bases at least Q20
Insert size: 169000; agarose-fp
Insert size: 161074; sum-of-contigs
Quality coverage: 3.60 in Q20 bases; agarose-fp
Quality coverage: 3.92 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
1
1124: contig of 1124 bp in length
1125: gap of unknown length
1225: contig of 1808 bp in length
3032: gap of unknown length
3132: contig of 1361 bp in length
4493: gap of unknown length
4594: contig of 1267 bp in length
5860: gap of unknown length
5960: contig of 1710 bp in length
7670: gap of unknown length
7671: contig of 1899 bp in length
9669: gap of unknown length
9770: contig of 1946 bp in length
11715: gap of unknown length
11815: gap of unknown length
14243: contig of 2428 bp in length
14244: gap of unknown length

16887: contig of 2544 bp in length
16987: gap of unknown length
19247: contig of 2260 bp in length
19347: gap of unknown length
21375: contig of 2028 bp in length
21475: gap of unknown length
25025: contig of 3550 bp in length
25125: gap of unknown length
28051: contig of 2926 bp in length
28151: gap of unknown length
32054: contig of 3903 bp in length
32154: gap of unknown length
35716: contig of 3562 bp in length
35816: gap of unknown length
40991: contig of 5075 bp in length
40991: gap of unknown length
44027: contig of 3036 bp in length
44127: gap of unknown length
48899: contig of 4772 bp in length
48900: gap of unknown length
54255: contig of 5256 bp in length
54355: gap of unknown length
58994: contig of 4639 bp in length
59094: gap of unknown length
64390: contig of 5296 bp in length
64491: gap of unknown length
70865: contig of 6375 bp in length
70865: gap of unknown length
78667: contig of 7702 bp in length
78667: gap of unknown length
87207: contig of 8440 bp in length
87307: gap of unknown length
97858: contig of 10551 bp in length
97859: gap of unknown length
109520: contig of 11562 bp in length
109521: gap of unknown length
120908: contig of 11288 bp in length
121009: gap of unknown length
141477: contig of 20469 bp in length
141478: gap of unknown length
162672: contig of 21095 bp in length
162772: gap of unknown length
163974: contig of 1202 bp in length
163975: gap of unknown length
164075: contig of 692 bp in length.

FEATURES             Location/Qualifiers
     source            1..164766
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="1"
                        /clone="RP11-245P10"
     misc_feature      1..1124
                        /note="assembly_name:Contig14"
     misc_feature      1225..3032
                        /note="assembly_name:Contig21"
     misc_feature      3132..4493
                        /note="assembly_name:Contig23"
     misc_feature      4594..5860
                        /note="assembly_name:Contig24"
     misc_feature      5961..7670
                        /note="assembly_name:Contig25"
     misc_feature      7771..9669
                        /note="assembly_name:Contig26"
     misc_feature      9770..11715
                        /note="assembly_name:Contig27"
     misc_feature      11816..14243
                        /note="assembly_name:Contig28"
     misc_feature      14344..16887
                        /note="assembly_name:Contig29"
     misc_feature      16988..19247
                        /note="assembly_name:Contig30"
     misc_feature      19348..21375
```



```

----- Genome Center -----
Center: Washington University Genome Sequencing C
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0661B12
----- Summary Statistics -----
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer ET; 100% of reads
Chemistry: Dye-terminator Big Dye; % of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 163945 bases at least Q40
Consensus quality: 167601 bases at least Q30
Consensus quality: 159687 bases at least Q20
Insert size: 148000; agarose-fp
Insert size: 173412; sum-of-contigs
Quality coverage: 5.67 in Q20 bases; agarose-fp

```


Quality coverage: 5.55 in Q20 bases; sum-of-contigs

- * NOTE: This is a 'working draft' sequence. It currently consists of 13 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
- * This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

- * 1 2478: contig of 2478 bp in length
- * 2479: gap of unknown length
- * 2579: contig of 4370 bp in length
- * 6948: gap of unknown length
- * 7049: contig of 2325 bp in length
- * 9373: gap of unknown length
- * 9474: contig of 3830 bp in length
- * 13303: contig of 4867 bp in length
- * 13403: gap of unknown length
- * 13404: contig of 4867 bp in length
- * 18270: gap of unknown length
- * 18370: contig of 7219 bp in length
- * 25589: gap of unknown length
- * 25690: contig of 9126 bp in length
- * 34815: gap of unknown length
- * 34916: contig of 10827 bp in length
- * 45743: gap of unknown length
- * 45843: contig of 17145 bp in length
- * 62988: gap of unknown length
- * 63087: contig of 24497 bp in length
- * 63088: gap of unknown length
- * 87585: contig of 27081 bp in length
- * 87685: gap of unknown length
- * 114766: contig of 27469 bp in length
- * 114866: gap of unknown length
- * 142334: contig of 32178 bp in length
- * 142335: Location/Qualifiers

FEATURES

- source 1..174612
 - /organism="Homo sapiens"
 - /mol_type="genomic DNA"
 - /db_xref="taxon:9606"
 - /chromosome="1"
 - /clone="RP11-661B12"
- misc_feature 1..2478
 - /note="assembly_name:Contig4"
- misc_feature 2579..6948
 - /note="assembly_name:Contig5"
- misc_feature 7049..9373
 - /note="assembly_name:Contig6"
- misc_feature 9474..13303
 - /note="assembly_name:Contig7"
- misc_feature 13404..18270
 - /note="assembly_name:Contig8"
- misc_feature 18371..25589
 - /note="assembly_name:Contig9"
- misc_feature 25690..34815
 - /note="assembly_name:Contig10"
- misc_feature 34916..45742
 - /note="assembly_name:Contig11"
- misc_feature 45843..62987
 - /note="assembly_name:Contig12"
- misc_feature 63088..87584
 - /note="assembly_name:Contig13"
- misc_feature 87685..114765
 - /note="assembly_name:Contig14"
- misc_feature 114866..142334
 - /note="assembly_name:Contig15"
- misc_feature 142435..174612
 - /note="assembly_name:Contig16"

ORIGIN

Query Match

32.1%; Score 321.6; DB 2; Length 174612;

Best Local Similarity 65.9%; Pred. No. 1e-58;
Matches 547; Conservative 0; Mismatches 244; Indels 39; Gaps 4;
QY 175 GCTGTGCTGCCAGGTTTCATAGAAAGATGTGAAAGACAGAGGCGCAGAGAGGGCCACG 234
Db 174102 GCCCTGCTGCCAGTTTCATCGGAAACTGAGAAACAGAGGCGCCACAGAGGGG-CACG 174044
QY 235 GCTGTGCTGCAAGTGTGAGCTGGAACAGTGCAGCCCTGTGAGTGGAGAAAGGGGTCTGAG 294
Db 174043 GCCACGCTGGCGGTGTGAGCTGAGCAAGACAGCCCTGTGAGTGGAGAAAGGGGTCCGAG 173984
QY 295 ACCCTCAGAGATGGGACAGATACAGCTCAGGCGAGGACCGGACTAAATGTGAGCTCCAG 354
Db 173983 ACCCTCAGAGATGGGACAGATCTGTCTGAGGACAGACGGGGCCATGTGTGAGCTCCAG 173924
QY 355 ATTGCTGGCTGGCCATGGCAGACACTGGGGAGTACTCTGTCGCTGTCGGGCGCAGGAGG 414
Db 173923 ATCCGTGGCTGGCCATGGTGGATGCGCGAGTACTCTGTCGTGTGTGAGAGGAGG 173864
QY 415 ACCTCGCTATGCTCACCGTCAGGCTCTACCCATCAAGTTTCACAGAGGGTCTTGAGGAAC 474
Db 173863 ACCTCAGCTCACTCACCATCAGCGTAAGACCGTGTATCCAGAGCGGTGTCCGGTGTCC 173804
QY 475 GA-----AGAGGCCACAGAGGGGCAACAGCC 501
Db 173803 AATTTTTTACCTCTGCTGTCTAGTACCCATTTCTAGTCTTACCTGACTGTGCGCATGCC 173744
QY 502 GTGCTGGCGTGTGAGCTGAGCAAGATGGCCCGCTGAGTGTGTGGAAGGGGCGCATGAGACC 561
Db 173743 TTCCAGCACATCACGAAATCCCTTCCAGCTTTTCTGGGCTGTGGAAGGAACACAGAGAAC 173684
QY 562 CTGAGAGATGGAGACAGACACAGCCCTGAGCAGAGCGGCGCAGGCTGAGCTGAGATC 621
Db 173683 CCGGCATTTGACCCAGACTCCACCCTCGAGCACCTCTGGGACCTGTCTCTGTGGGAGTC 173624
QY 622 CGCGGCTCTGTGCGCAGAGGACGCTGGGGAGTACTCTGTGATGTGCGGAGGAGGAGACC 681
Db 173623 TGTCTTATCCT-CCGCTTGGCATGCGCTGTCCCTGTGCTTCCAGAACTCTGATCTCC 173565
QY 682 TCAGCCA----TGCTCACCGTCAGGCGCATGCTTCAAGTTTCATAGAGGCTTGAGGAA 737
Db 173564 ATGCTGTGTCTCCATCTCTCCCGACCATGCTGCCCACTTCATAGGAAGACTTGAGACA 173505
QY 738 TGAAGAGGCCACAGAAAGGGGACACGCGCCACGCTGTGGTGTGAGCTGAGCAAGCGGCGACC 797
Db 173504 CCAAGAGAGCATAGAAAGGGGCGCACAGCCAGCTGCGGTGTGAGCTGAGCAGGCGGCC 173445
QY 798 GGTGGAGTGGAGAGGGGCGCATGAGACCTCTAGAGATGGGGACAGACACAGCCTTGAGGCA 857
Db 173444 CGTGGAGTGGAGAGGGGCGCTGAGAGCCTCAGAGATGGGGACAGACATAGCCTTGAGGCA 173385
QY 858 GGAACGGGTCCAGTGTGAGCTGCAGATCGTGGCTGCGCTGTGGTGTGGTATGCCGGGAGTA 917
Db 173384 GGACGGGCTGTGTGCGAGCTGCAGATCTGTGGCTGTGGCTGTGGCAGATGTGGGGAGTA 173325
QY 918 CTCGTGTGTGTGCGGCGCAGGAGGACCTTCAGCCACACTCACTGTGCGAGG 967
Db 173324 CTCCTGTGTGTGCGGAGGAGGAGGACCTCTGCCACTCTCACCGTGAAG 173275

RESULT 9

AC026657/c

LOCUS

DEFINITION

SEQUENCE, 31 unordered pieces.

AC026657

AC026657.4

GI:9958202

HTG; HTGS_PHASE1; HTGS_DRAFT.

HTG; HTGS_PHASE1; HTGS_DRAFT.

HTG; HTGS_PHASE1; HTGS_DRAFT.

HTG; HTGS_PHASE1; HTGS_DRAFT.

HTG; HTGS_PHASE1; HTGS_DRAFT.

HTG; HTGS_PHASE1; HTGS_DRAFT.

HTG; HTGS_PHASE1; HTGS_DRAFT.

HTG; HTGS_PHASE1; HTGS_DRAFT.

HTG; HTGS_PHASE1; HTGS_DRAFT.

HTG; HTGS_PHASE1; HTGS_DRAFT.

misc_feature	/note="assembly_name:Contig40"			
	59095..	.64390		
misc_feature	/note="assembly_name:Contig41"			
	64491..	.70865		
misc_feature	/note="assembly_name:Contig42"			
	70966..	.78667		
misc_feature	/note="assembly_name:Contig43"			
	78768..	.87207		
misc_feature	/note="assembly_name:Contig44"			
	87308..	.97858		
misc_feature	/note="assembly_name:Contig45"			
	97959..	.109520		
misc_feature	/note="assembly_name:Contig46"			
	109621..	.120908		
misc_feature	/note="assembly_name:Contig47"			
	121009..	.141477		
misc_feature	/note="assembly_name:Contig48"			
	141578..	.162672		
misc_feature	/note="assembly_name:Contig49"			
	162773..	.163974		
misc_feature	/note="assembly_name:Contig9"			
	164075..	.164766		
ORIGIN				
Query Match	27.8%; Score 278.6; DB 2; Length 164766;			
Best Local Similarity	62.6%; Pred. No. 1.5e-49;			
Matches 559; Conservative	0; Mismatches 234; Indels 100; Gaps 4;			
QY	175	GCTCTGCCCTGCCAGGTT	CATAGAAGATCTGA	AAACAGGAGGCCACAGAGGGGCCACG 234
Db	133606	GCCTCCGCCGAGATT	CATAGAAGATGTGA	AAATCACGAGGCCACAGAGGGGGCCACA 133547
QY	235	GCTGTGCTGCAGTGT	GAGTGAACAGTGC	AGTGCAGCCCTGTGGAGTGGAGAAAGGGGTCTGAG 294
Db	133546	GCTGTGCTGCAGTGT	GAGTGAACAGTGC	AGTGCAGCCCTGTGGAGTGGAGAAAGGGGTCTGAG 133487
QY	295	ACCTCAGAGATGGGA	CAGATACAGCTTA	AGCGGACGGGACTAAATGTGAGCTGCAG 354
Db	133486	ACCTCAGAGATGGGA	CAGATATAGCTTA	AGCGGACGGGACTAAATGTGAGCTGCAG 133427
QY	355	ATTGTGGGCTGGCC	ATGCGACACTGGG	AGTACTCGTCCGTGTCGGGCGAGGAGG 414
Db	133426	ATTGTGGCTGGCT	GTGGAGACACTGG	AGATTTGTGTGTGTCGGGCGAGGAGAG 133367
QY	415	ACCTCGGCTATGCT	CACCTCAGGCTCT	ACCCATCAAGTTTCACAGAGGGTCTTGAGGAAC 474
Db	133366	ACCTCAGCTTACCT	CACCTGTCAGGGT	AAAGATCCTGTGTGGTCAATGACTTGTGGCT 133307
QY	475	GAAGAGGCCACAG	AGGGGGCAACA	-----GCGTGTGCGGTTGT 513
Db	133306	TGGAAATCTAAGT	GTCTGTGTCTAC	CCACTTCTACTTCCAAATGTGGACCCCTGTGGGA 133247
QY	514	GAGCTGAGCAGAT	GGCCCGCTGGAG	TGTTGGAGGGGCAT-----GAGACCTCTCAG 566
Db	133246	ACCTTGAGGTGT	GTGTCTCTTAAGT	GGGTCTTCTGGGAGTCTCTGTCGCTCTGCTTC 133187
QY	567	AGATGGAGACAGA	CACAGCTGAGG	CAGGACGGGGCCAGGTGTGAGCTGCAGATCCGCGG 626
Db	133186	CTCATCAGGAGAT	GTCTCTCTGTTT	CATCCCATCCCTGCGTGTCTCTGTCTCTCTTGGG 133127
QY	627	CCTCGTGGCAGAG	CGCTGGGA	----- 650
Db	133126	CCATGTGCTTTCT	CGCCTTCGGTAC	ATTCATTTGCTTATTTCTGCCATGGTCTCTTGGTGG 133067
QY	651	-----	GTACCTGTGCAT	GTGCGGAGGAGGAGGACCTCAGCCATGCTCA- 694
Db	133066	TTTGTGTGGCTCT	CGTGTGTCTGTGT	CCTGTCCACAGAGTTGTCCACAGTCTGAT 133007
QY	695	-----	CCGTCAGGGCCA	TGCCTTCCAGTTTCATAGAGGGTCTGAG 734
Db	133006	GATCTCAGTACTG	CTTGGTCTTCC	CAGCGCTTCCAGCCAGATTCATAGACAAATGAC 132947

QY	735	GAATGAAGAGCCACAGAGGGGACACGGCCACGCTGTGTGTGCTGAGTGCACAGGGCCG	794
Db	132946	AAACAGGAGCCACAGAGAGGGGCCACGGCCACACTGCCTGTGAATGAGCAAGGTGGC	132887
QY	795	ACCGTGTGAGTGGAGGAGGGGCATGAGACCTCAGAGATGGGACACAGACACACCTTGAG	854
Db	132886	CCCTGTGAGTGGAGGAGGGACCTGAAACCTCCGAGATGGGACACAGACACCTTGAG	132827
QY	855	GCAGGACGGTTCAGGTGTGAGCTGCAGATCCGTGGCTGGCTGTGTGTGATGCCGGGA	914
Db	132826	GCAGGATGGGACCAAGGTGTGAGCTGCAGATTCGTGGCTGTCTGTGGCAGATGCCGGGA	132767
QY	915	GTACTCTGTGTGTGCGGGCAGGAGAGGACCTCAGGCACACTCACTCTCAGGG	967
Db	132766	GTACTCTGTGTGTGCGGGCAGGAGAGGACCTCAGGCACACTCAGCATCAGGG	132714
RESULT 10			
AC099089/c			
LOCUS			
Rattus norvegicus clone CH230-154E3, WORKING DRAFT SEQUENCE.			
AC099089			
AC099089.5 GI:30522701			
HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.			
Rattus norvegicus (Norway rat)			
Rattus norvegicus			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;			
Rattus.			
1 (bases 1 to 260998)			
REFERENCE			
AUTHORS			
Muzny,D.,Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,			
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,			
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,			
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,			
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,			
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderson,E.,			
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,			
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,			
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,			
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,			
Delgado,O., Denison,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,			
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,			
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,			
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,			
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,			
Gebregorgis,E., Geer,K., Gill,R., Grady,M., Guerri,W., Guevara,W.,			
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,J.,			
Harvey,X., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,			
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogue,M.,			
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,			
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,			
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,			
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,			
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,			
Lorenshuwa,L., Loundsege,H., Lozado,R.J., Lu,X., Ma,J.,			
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,			
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,			
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,			
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Naif,L.,			
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,			
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,			
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,			
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.,			
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,			
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,			
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,			
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatman,S., Shen,H.,			
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajls,D.,			
Sheddy,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,			
Steinle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,			
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Umani,K.,			
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,			

Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 260998)
Worley,K.C.
Direct Submission
Submitted (09-NOV-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 260998)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On May 10, 2003 this sequence version replaced gi:22855456.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GGLZ
Center clone name: CH230-154E3
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 216929 bases at least Q40
Consensus quality: 219228 bases at least Q30
Consensus quality: 220845 bases at least Q20
Estimated insert size: 226118; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html.)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 260998: contig of 260998 bp in length.
Location/Qualifiers
1. .260998
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-154E3"
1. .1486
misc_feature
1. .1486
/note="wgs_contig"

FEATURES
source
misc_feature
ORIGIN
Query Match 27.6%; Score 276; DB 2; Length 260998;
Best Local Similarity 61.1%; Pred. No. 5e-49;

Direct Submission
 Submitted (04-APR-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Apr 7, 2002 this sequence version replaced gi:19699553.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-192P17 is from the RPCI-23 Mouse PAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
 VECTOR: pBACE3.6.

FEATURES

Location/Qualifiers

1..211829
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="11"
 /clone="RP23-192P17"
 /clone_lib="RPCI-23"

ORIGIN

Query Match 26.7%; Score 267.2; DB 10; Length 211829;
 Best Local Similarity 60.7%; Pred. No. 3.9e-47;
 Matches 477; Conservative 0; Mismatches 299; Indels 11; Gaps 2;
 QY 182 CTGCGAGTTCATAGACATGTGAAACACAGGAGGCGGACGAGAGGGCCACGCTGTC 241
 Db 51183 CTGCGAGTTCATAGAGAACTGAGAAACAGGAGGCGGACGAGAGGACACGCTCAC 51124
 QY 242 TGCAGTGTAGCTGAACAGTGCAGCCCTGTGGAGTGGAGAAAGGGTCTGAGACCCCTCA 301
 Db 51123 TTCGGTGTGAGCTGACCAAGAGGCCCCCGTGGAGTGGAGAGGAACAGAGACCTGA 51064
 QY 302 GAGATGGGGACAGATACAGCTTGAGGCGAGGACGGGACTAATGTGAGCTGCAGATTCGTG 361
 Db 51063 GAAATGGGGACAAATACAGTCTGAAGCAGGATGGAGCTGTGTGTGAGCTGCAGATCTGTA 51004
 QY 362 GCCTGGCCATGGCAGACACTGGGAGTACTCGTGGTGTGGGCGAGGAGGACCTCGG 421
 Db 51003 GTCTGCTGGTGCAGATGCAGGGAGTACTCGTGTGTGTGGGCGAGGAGACTTCAG 50944
 QY 422 CTATGCTCACCGTCAGGGCTCTACCCATCAAGTTCAAGAGGGTCTGAGGAACGAAGAGG 481
 Db 50943 CCAGCGTACTGTCAAGGGT-----AAAGAGCGTGCATGGCTTCCCTCCCCCGCTGT 50888
 QY 482 CCACAGAGGGGCAACAGCGTGTGGGCTGTGAGCTGAGCAGATGGCCCGCTGGAGT 541
 Db 50887 GTACTACTGTGCAATCCACCTTACTTCTTGAGTCATTATATATATGAAACGAGAAAT 50828
 QY 542 GGTGGAAGGGGCATGAGACCTTCAGAGATGGAGACACACAGCCTTGAGGACGACGGGG 601
 Db 50827 ACCAGCATCTGCAGCTGGCTTCCACTTTCAGTGTGGTGGGTCTCTCTCTGTGCTGC 50768
 QY 602 CCAGGTGTGAGCTGCAGATCGGGGCTCTGGGCGAGGAGCGTGGGGAGTACTCTGTGCA 661
 Db 50767 ATTCTATCTTGTCCCTATCC-----TGTTCTATGTCACTTCTATCTTCCCACTC 50715

QY 662 TGTGGGGAAGAGAGGACCTCAGCCATGCTCACCCTCAGGGCCATGCTTCCAGTTCA 721
 Db 50714 TGTAGCAGATCTGATGCTCCAATCTTTGTCTTACCTCTAGCTCTGCTTCCACTTG 50655
 QY 722 TAGAGGGTCTCAGGAATGAAGAGGCCACAGAAAGGGGACAGGCCACGCTGTGTGAGC 781
 Db 50654 TCAGNAGGCTCAGAGGACNAGAGGCCACAGAAAGAGACACGACCACTTGCAGTGTGAAC 50595
 QY 782 TGAGCAAGGGGCAACCGTGGAGTGGAGGAGGGGCATGAGACCTCAGAGATGGGGACA 841
 Db 50594 TGAGCAAGGGTGGCCCGTGGAGTGGAGGAGGGAACAGAGACCTCAGAGATGGGGACA 50535
 QY 842 GACACAGCTGAGGACGAGGCTCCAGGTGTGAGCTGAGATCCGTGGCTGTGGTGG 901
 Db 50534 GATCACCTGAAACAGGATGGGGCTGTGTGTGAGCTGCAAAATCCCGAGGCTGACCATAG 50475
 QY 902 TGGATGCCGGGGAGTACTCGTGTGTGCGGGGAGGAGAGGACCTCAGCCACACTCACTG 961
 Db 50474 CAGATGCTGGCGAGTACTTGTGCACGTTGTGACAGGAAAGACCTCAGCCACGCTGACC 50415
 QY 962 TCAGGG 967
 Db 50414 TCAGGG 50409

RESULT 12

AX463538
 LOCUS AX463538 2155 bp DNA linear PAT 15-JUL-2002
 DEFINITION Sequence 102 from Patent WO248337.
 ACCESSION AX463538
 VERSION AX463538.1 GI:21886312
 KEYWORDS Homo sapiens (human)
 SOURCE Homo sapiens
 ORGANISM Homo sapiens
 REFERENCE 1
 AUTHORS Duggan,B.M., Yao,M.G. and Griffin,J.A.
 TITLE Secreted proteins
 JOURNAL Patent: WO 0248337-A 102 20-JUN-2002;
 INCYTE GENOMICS INC (US)
 FEATURES
 source
 1..2155
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 2902793CB1"
 ORIGIN

Query Match 23.9%; Score 238.8; DB 6; Length 2155;
 Best Local Similarity 59.8%; Pred. No. 1e-40;
 Matches 531; Conservative 0; Mismatches 262; Indels 95; Gaps 4;
 QY 175 GCTCTGCTGCCAGGTTTCATAGAGATGTGAAAAACAGGAGCCAGAGAGGGGCCACG 234
 Db 205 GCGCTGCTGCCAGGTTTCAGAGGGTTCAGAGAAATGAAGAGGCCATGAAGAGGGGCCACA 264
 QY 235 GCTGTGCTGCAGTGTGAGCTGAACAGTGCAGCCCTGTGGAGTGGAGAAAGGGGTCTGAG 294
 Db 265 GCACACTGCAATGTAGCTGAGCAAGGAGCCCTCTGGAGTGGAGAAAGGGCCCTTGAG 324
 QY 295 ACCCTCAGAGATGGGGACAGATACAGCCCTGAGGACGAGGACGGGACTAAATGTGAGCTGAG 354
 Db 325 GCTCTCAGAGATGGGGACAAATACAGCTGAGACAAGACGGGGCTGTGTGAGCTGAG 384
 QY 355 ATTCTGCGCTGGCCATGGCAGACACTGGGGAGTACTCTGCTGTGTGCGGGCAGAGAGG 414
 Db 385 ATTCTGCGCTGGCTATGGCAGATAACGGGGTGTACTCATGTGTGTGGGCGAGAGAGG 444
 QY 415 ACCTCGGCTATGCTCACCGTCAAGGGCTCTACCCATCAAGTTTCACAGAGGGTC----- 466
 Db 445 ACCTCAGCTACACTCCTGTGAGGGGTAAGATCCTATGTGGCCATGTGGGCTTGTGGCT 504


```
QY 467 TGAGAACGAGAGAGCCACAGAGAGGGGCAACAGCGTGTCTGCGGTGTGAGCTGAGCAAGA 526
Db 505 TGGTGATATACACCTCTCTGTGTGTCACCACTTCTGCCTCCAAATGGGCACATCTCTGTG 564
QY 527 TGGCCCCGTGGAGTGTGGTGGAGGGGCGATGAGACCTCAGAG-----568
Db 565 GAAACCTGTGATTGTCTGTCTCTTAACTGGGGCCCTCTAGGTATCTCTGCTGCCCTCCCTT 624
QY 569 --ATGGAGACAGACAGCCTGTGAGCAGGAGCGGGGCCAGGTGTGAGCTGCAGATCCGGCG 626
Db 625 CTCATCAGTAGATGTCTCTCTCTCATGATGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 684
QY 627 CCTGTGCGCAGAGGAGCGTGGGA-----650
Db 685 CCAATTGGCTTCTCATTTGTTATATATTGCTCATCTAGCTATGGTCCCTTTGGTGGTTTGT 744
QY 651 -----GTACCTGTGCATGTGCGGGAAGGAGGAGGACCTCAGCCATGCTCA-----694
Db 745 GTGGCTCCTATTGTGTTCATGTTCTGTCCGAAAATCCTCCAGACAGTCTGATGATAT 804
QY 695 -----CCGTGAGGCGCATGCTTCCAAAGTTTCATAGAGGCTCTGAGGAATG 739
Db 805 CAGTGACTGTTTGGTCTTCTCAGCCCTGCCTGCCAGATTTCATAGAGGATATGAGAAACC 864
QY 740 AAGAGCCACAGAGGAGGACAGCGCCAGCTGTGTGTGAGCTGAGCAAGCGGCACCGG 799
Db 865 AGAAGGCCACAGAGAGGGGCTCAGTCAATGTAAGCTGAGAAAGCGCGCCCGC 924
QY 800 TGGAGTGAGGAGGCGCATGAGACCTCAGAGATGGGACAGACACAGCCTGAGGCGAGG 859
Db 925 TGGAGTGAGGAGGCGCCACACCTCAAGATGGGACAGGTACAGCCTGAGGCGAGG 984
QY 860 ACGGTTCAGGTGTGAGCTGAGATCCGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 919
Db 985 ATGGACACAGTGTGAGCTGAGATTCGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 1044
QY 920 CGTGTGTGTGGGCGAGAGGAGGACCTCAGCCACACTCACTGTACGG 967
Db 1045 CGTGATATGTGAGCAGGAGGAGGACCTCGGCCACGCTCACTGTACGG 1092

RESULT 13
LOCUS HSM800639 HSM800639 2295 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFPZ434C153 (from clone DKFPZ434C153);
partial cds.
ACCESSION AL080128
VERSION AL080128.1 GI:5262565
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 2295)
JOURNAL Wambutt,R., Heubner,D., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
Direct Submission
Submitted (16-JUN-1999) MIPS, Am Klopferstritz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFPZ434C153) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
*Location/Qualifiers
1..2295
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFPZ434C153"

FEATURES
source
```

```
/tissue_type="testis"
/clone_lib="434 (synonym: htes3) Vector pSport1; host
DH10B; sites NotI + SalI"
/dev_stage="adult"
1..2295
/gene="DKFPZ434C153"
<1..348
/gene="DKFPZ434C153"
/notes="similarity to titin"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB45726_1"
/db_xref="GI:5262566"
/db_xref="UniProt/TREMBL:Q9Y4R5"
/translation="AMFCPKNPDPDSLMISVTWVSFSAIPARPIEDMRNOKATEGATVT
LQCKLRKAAPVEMRWKGNPTLKDGRYSYKQDGTSCELQIRGLVIADAGEYSCICRQER
TSATLTVRGKDHM"
2232..2237
/polyA_signal /gene="DKFPZ434C153"
2262
/polyA_site /gene="DKFPZ434C153"

ORIGIN
Query Match 19.4%; Score 193.8; DB 9; Length 2295;
Best Local Similarity 78.8%; Pred. No. 4e-31; Mismatches 0; Gaps 0;
Matches 231; Conservative 0; Indels 62;

QY 167 CCATCAGGCGCTCTGCTGCCAGGTTTCATAGAGATGTGAAAAACCCAGGAGCCAGAGAAG 226
Db 59 CTTCTCAGCCTGCTGCCAGATTTCATAGAGATATGAGAAACCCAGGAGCCAGAGAAG 118
QY 227 GGGCCACGGCTGTGTCAGTGTGAGTGTGACAGTGTGAGCCCTGTGAGTGGAGAAAGG 286
Db 119 GGGCTACAGTACATTCGAATGTAAGCTGAGAAAGCGGCCCTGTGAGTGGAGAAAGG 178
QY 287 GGTCTGAGACCTCAGAGTGGGACAGATACAGCTGAGCAGCAGGAGCCGGAATAATGTG 346
Db 179 GGCCCAACACCTCAAGATGGGACAGGTACAGCTCAAGCAGATGGGACAGTGGACAGTGTG 238
QY 347 AGCTGCAGATTCGTGGCCCTGGCCATGGCAGACACTGGGGAGTACTCTGCTGCTGCGGCG 406
Db 239 AGCTGCAGATTCGTGGCCCTGTCTATAGCAGATCTGAGAGATATCTCTGTCATATGTGAGC 298
QY 407 AGGAGAGACCTCGGCTATGTCTACCGCTCAGGGCTTACCCATCAAGTTTCA 459
Db 299 AGGAGAGACCTCGGCCACGCTCACTGTGAGGGGTAAAGACCCACATGTGACCA 351

RESULT 14
LOCUS BV014682/c BV014682 621 bp DNA linear STS 30-MAY-2003
DEFINITION S212P60282FH4.T0 CZECHII/BI Mus musculus STS genomic, sequence
tagged site.
ACCESSION BV014682
VERSION BV014682.1 GI:31098577
KEYWORDS STS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE 1 (bases 1 to 621)
JOURNAL Wade,C.M., Kulbokas,E.J. III, Kirby,A.W., Zody,M.C., Mullikin,J.C.,
Lander,E.S., Lindblad-Toh,K. and Daly,M.J.
MEDLINE The mosaic structure of variation in the laboratory mouse genome
PUBMED Nature 420 (6915), 574-578 (2002)
COMMENT 22354684
12466852
Contact: Kerstin Lindblad-Toh
Whitehead Institute for Biomedical Research, Center for Genome
Research
320 Charles Street, Cambridge, MA 02141, USA
Tel: 6172521477
```


Fax: 6172580903
 Email: kereli@genome.wi.mit.edu
 Primer A: None
 Primer B: None
 STS size: 621
 Protocol:
 WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SVIMJ, C3H/HeJ, and BALB/cBYJ. The WGS reads were placed uniquely on the MGSv3 C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated
 as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J and the strain from which the particular read came. The validation rate for these SNPs was estimated at approximately 98%.
 Location/Qualifiers
 1. .621
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /strain="CZECHII/EI"
 /db_xref="taxon:10090"
 /map="11 22-567 59585524-59586069"
 /clone_lib="CZECHII/EI"
 <1. .>621

FEATURES
 source
 Query Match 18.5%; Score 185; DB 11; Length 621;
 Best Local Similarity 81.1%; Pred. No. 3.8e-29;
 Matches 215; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

STS
 ORIGIN
 Query Match 18.5%; Score 185; DB 11; Length 621;
 Best Local Similarity 81.1%; Pred. No. 3.8e-29;
 Matches 215; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 703 GCCATGCGCTTCCAGTTCATAGAGGCTCTGAGGATGAGAGGCCACAGAGGGGACACG 762
 Db 477 GCCCTGCCCCCTAGATTATGAAAGATTGAGAGCCAGAGGCCAGGAGGACTGTG 418
 QY 763 GCCACGCTGTGTGTGAGTGAAGCAAGCGGCAACCGGTGAGTGAAGAGCGGCATGAG 822
 Db 417 GCCACTCTGAGGTGCAGATGAGCAAGGCTGCCCACTGAGTGAAGAGGGTTCAGAA 358
 QY 823 ACCCTCAGAGTGGGACAGACAGCTGAGGAGGAGGACCGGTCCAGGTGTGAGTCCAG 882
 Db 357 ACCCTGAAAGATGGGACAGATACAGCTGAGGAGGAGTGGGGCCATGTGTGAGTCCAG 298
 QY 883 ATCCGTGCGCTGCTGTGTGTGATCCGGGAGTACTCTGTGTGTGGGCGAGGAGG 942
 Db 297 ATCCGTGACCTGGCTGTGTGAAGTACTGGGAGTACTCATGTGTGTGGGCGAGGAG 238
 QY 943 ACCTCAGCCACACTCACTGTCAAGG 967
 Db 237 ACCTCAGCCACACTGAGTGTCAAGG 213

RESULT 15
 AX834822 2534 bp DNA linear PAT 15-DEC-2003
 LOCUS
 DEFINITION Sequence 1946 from Patent EP1347046.
 ACCESSION AX834822
 VERSION AX834822.1 GI:39920957
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 AUTHORS Isogai,T., Sugiyama,T., Otsuki,T., Wakamatsu,A., Sato,H., Ishii,S., Yamamoto,J.I., Isono,Y., Hio,Y., Otsuka,K., Nagai,K., Irie,R., Tamechika,I., Seki,N., Yoshikawa,T., Otsuka,M., Nagahari,K. and Masuho,Y.
 Full-length cDNA sequences
 Patent: EP 1347046-A 1946 24-SEP-2003;
 Research Association for Biotechnology (JP)
 Location/Qualifiers
 1. .2534
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"

/db_xref="taxon:9606"
ORIGIN
 Query Match 10.7%; Score 107.4; DB 6; Length 2534;
 Best Local Similarity 47.7%; Pred. No. 1e-12;
 Matches 420; Conservative 0; Mismatches 446; Indels 15; Gaps 3;

QY 136 TCGCGGAGGAGAGAGAGCTCAGCCACGCTACCATCAGGGCTCTGCTGCAGATTATA 195
 Db 199 TCCCTGAAGGTTCGACCTCTGCCCGCTCGAGGTCGAGAGCGGCTGTTCTCTG 258
 QY 196 GAAGATGTGAAAAACCAAGAGGAGCCAGAGAGGGGCCACGGCTGTGCTGCAGTGTGAGCT- 254
 Db 259 AAGCGCTGGATGACCTGTCGAGAGAGGAGCGCGGACCTTGGCCCTCGAGTGTGAAGTC 318
 QY 255 --GAACAGTCAGCCCTGTGAGTGTGAGAAAGGGGTCTGAGACCTCAGAGATGGGGAC 312
 Db 319 TCTGACCCGAGGCCCATGTGTGTGGCGCAAGATGGCGTGCAGCTGGGCCCCAGTGAC 378
 QY 313 AGATACAGCCTGAGGAGGAGCGGACTAAATGTGAGCTGCAGATTCTGTGGCCTGGCCATG 372
 Db 379 AAGTATGACTTCTGACACAGCGCGGGCAACCGGGGGCTCTGTGTGCATGACGTGAGCCCT 438
 QY 373 GCAGACACTGGGAGTACTCTGTGTGTGGGGCAGGAGGAGACCTCGGCTATGCTCACC 432
 Db 439 GAAGAGCGCGGCTGTACACTTGCACGCTGGGCTCCGAGGAGACCCGGGCCCGGTCCGC 498
 QY 433 GTCAAGGCTTACCCATCAAGTTCACAGAGGCTCTGAGGAACGAAAGAGGCCACAGAGGG 492
 Db 499 GTGCACGATTCGACGCTGGGCATCAACAGAGGCTGAAGACAATGAGGTGTCTGAAGGG 558
 QY 493 GCAACAGCCGTGTGCGGTGTGAGCTGAGCAAGATGCCCGCTGCGGAGTGTGGAGGGG 552
 Db 559 GAAAGCTGCAGCTTTGAGTGCCTCTGTCCACAGAGTGCAGGCCACCGGCCCATGTGG 618
 QY 553 CATGAGACCTCTAGAGATGAGACAGACAGCTTGAGGCAGGAC-----GGGGCC 603
 Db 619 ACAGTCGCTGGGAAGACAGTGGGCAGCTCCAGCGCTTCCAGGCCACACGTCAGGGCCGA 678
 QY 604 AGGTGTGAGTGCAGATCCGCGGCTCTGTGGCAGAGAGACGCTGGGAGTACCTGTGCATG 663
 Db 679 AATATCATCTGTGTGCTGGGAGGCTGCACCAAGTGTATCCCGGGGAGGTGTTCTCT 738
 QY 664 TCGCGGAAAGAGAGGAGACCTCAGCCATGCTCACCGTCAGGGCCATGCTTCCAGTTTATA 723
 Db 739 GTGCGGGGCTCACCTCCAGGCTTCACTCATTTGTTCAGAGAGGCGCGGCCCATCATC 798
 QY 724 GAGGCTCTGAGGAATGAAGAGGCCACAGAAAGGGGACACGCGCCACGCTGTGTGTGAGCTG 783
 Db 799 AAGCCCTTGAAGACCAAGTGGGTGGCGCCAGGGGAGGACGTGGAGCTGCGCTGTGAGCTG 858
 QY 784 AGCAAGGCGG---CACCGGTGGAGTGGAGAGAGGGGCGCATGAGACCTCAGAGATGGGGAC 840
 Db 859 TCACGGCGGGAGACGCGCCGTGCATCTGTGTGAAGGACAGAAAGGCCATCCGCAAGAGCCAG 918
 QY 841 AGACACAGCCTTGAGGAGGAGCGGCTCCAGGTGTGAGCTGCAGATCCGTCGCTGGCTGTG 900
 Db 919 AAGTATGATGTGTCTGCGAGGGGACGATGGCCATGCTGTCTATCCGCGGGGCTCGCTC 978
 QY 901 GTGATGCGCGGGAGTACTCTGTGTGTGTCGGGCGAGGAGGAGGACTTCAGCCACTCACT 960
 Db 979 AAGGACGCGCGGGGAGTACACGCTGTGAGGTGAGGCTTCCAAAGAGCACAGCCAGCCCTCAT 1038
 QY 961 GTCAGGGCCCTGCTGCACGATTTATAGAAGATGTGAAAAA 1001
 Db 1039 GTGGAAGAAAAAGCAACTGCTTTCAGAGGAGTGCACCA 1079

Search completed: March 21, 2005, 05:34:49
 Job time : 4533.62 secs

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 16:50:48 ; Search time 3663.47 Seconds
(without alignments)
10400.615 Million cell updates/sec

Title: US-10-077-130-4_COPY_10500_11500

Perfect score: 1001
Sequence: 1 gcccttgagtgaggaa.....tccatagaagatgtgaaaaa 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_est3:*

4: gb_est4:*

5: gb_est5:*

6: gb_est6:*

7: gb_est7:*

8: gb_est8:*

9: gb_est9:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	377.4	37.7	595	2	BE086827 QV1-BT067
2	368.8	36.8	536	2	BF824937 ILO-HN003
3	340.6	34.0	486	4	BM030364 488988 MA
C 4	302	30.2	723	5	BQ446463 UI-H-EU1
C 5	278	27.8	487	2	BE466441 hz21h03.x
C 6	262.6	26.2	437	2	BF826423 CM4-HN002
7	218.2	21.8	369	2	BF826489 CM4-HN002
8	216	21.6	372	7	CV315157 CM1-HT087
C 9	196.2	19.6	570	9	CE836029 tigr-g88
C 10	194.4	19.4	359	2	BF853346 MR2-EN009
C 11	186	18.6	515	7	CV346932 MR2-BN038
C 12	175.2	17.5	601	2	BF398067 UI-R-B82
C 13	173.2	17.3	324	5	BQ361458 PMO-OT023
C 14	173.2	17.3	690	8	BH025901 RPCI-24-2
C 15	147	14.7	204	4	BG987697 PM3-HT116
C 16	137.4	13.7	566	7	CN664157 A0809E09-
17	135.6	13.5	505	9	CE078765 tigr-g88-
18	135	13.5	1074	4	BJ730664 BJ730664
C 19	131	13.1	247	4	BG987696 PM3-HT116
C 20	129.8	13.0	742	7	CV352920 MR3-HT107
C 21	125.8	12.6	743	5	BV358283 603477905
C 22	125	12.5	647	8	AZ590254 IM0399B18
23	121.4	12.1	657	8	AZ506477 IM0347H09
24	115	11.5	628	8	AZ985514 2M0267G22

C 25	112.4	11.2	350	2	BF748073
C 26	108.4	10.8	380	4	BF957683
C 27	87.6	8.8	212	2	BF827521
C 28	81	8.1	382	1	AI478779
C 29	81	8.1	740	5	BU689712
C 30	78.6	7.9	330	8	AQ237886
C 31	72.8	7.3	510	5	AX478907
C 32	71.2	7.1	173	2	BE086762
C 33	69.2	6.9	1010	9	CNS02HPP
C 34	61.6	6.2	584	7	CK829379
C 35	60.4	6.0	899	9	CL073463
C 36	60.2	6.0	1970	9	CG748837
C 37	60.2	6.0	5558	3	AK029863
C 38	60	6.0	997	9	CNS05GPF5
C 39	58.4	5.8	895	2	BE731455
C 40	57.8	5.8	586	9	BX153687
C 41	57.8	5.8	1132	5	BQ401116
C 42	57.2	5.7	1153	8	CC252982
C 43	57	5.7	332	8	AZ754571
C 44	57	5.7	1393	9	CL043635
C 45	56.4	5.6	1184	5	BM911719

ALIGNMENTS

RESULT 1
BE086827
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

BE086827 595 bp mRNA linear EST 12-JUN-2000
QV1-BT0678-300400-182-f06 BT0678 Homo sapiens cDNA, mRNA sequence.

BE086827.1 GI:8477221
EST.

Homosapiens (human)

Homosapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 595)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,

Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,

O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

20202663

10737800

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: asimpson@ludwig.org.br

This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2=QV1-BT0678-300

400-182-f06&t3=2000-04-30&t4=1)

Seq primer: puc 18 forward

High quality sequence start: 11

High quality sequence stop: 595.

Location/Qualifiers

1. 595

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/dev_stage="Adult"

/clone_lib="BT0678"

/note="Organ: breast; Vector: puc18; Site 1: Sma1; Site 2:

Sma1; A mini-library was made by cloning products derived

ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.
1 (bases 1 to 486)
Smith, T.P.L., Grosse, W.M., Freking, B.A., Roberts, A.J., Stone, R.T., Casas, E., Wray, J.E., White, J., Cho, J., Fahrenkrug, S.C., Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A., Chitko-McKown, C.G., Perce, G., Holt, I., Karamecheva, S., Liang, P., Quackenbush, J., and Keefe, J.W.
Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle
Genome Res. 11 (4), 626-630 (2001)
21180013
21182978
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt_trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -minmatch 12 options.
PCR Primers
FORWARD: AGGAACAGCTATGACAT
BACKWARD: GTTTCACGTCAGCG
Plate: 120 row: E column: 13
Seq primer: ATTAGTGACACTATAG.
Location/Qualifiers
1. 486
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 280V"
/note="Vector: PCMV SPORT6; Site 1: NotI; Site 2: SalI; Library made from pooled tissue from testis, thymus, semitendinosus muscle, longissimus muscle, pancreas, adrenal, and endometrium."

FEATURES

source

ORIGIN

Query Match 34.0%; Score 340.6; DB 4; Length 486;
Best Local Similarity 81.6%; Pred. No. 1e-68;
Matches 394; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

Qy 515 AGCTGAGCAAGATGTCGCCCGTGGAGTGGTGAAGGGCATGAGACCTCAGAGATGGAG 574
Db 1 AGCTGAGCAAGCGGCCCGTGGAGTGGAGTGGAGAGGGCCCGAGACCTTAGAGCTGGGG 60

Qy 575 ACAGACACAGCTGAGGCGAGGACGGGGCGAGTGTGAGCTGCAGATCCGCGCCTCGTGG 634
Db 61 ACAGACTCGGGTGAAGCAGGACGGGGCGGTGTGTGAGCTGGATCCATGACCTGACCA 120

Qy 635 CAGAGGACGCTGGGAGTACCTGTGCTGTCGCGGAAGAGGACCTCAGCCTGCTCA 694
Db 121 TGGAGGATGCTGGGAGTACTCATGATATGCGGCGAGAGAAAACCTCAGCAGCTCG 180

Qy 695 CCCTCAGGCGCATGCTTCCAAAGTTTCATAGAGGCTCAGGATGAAGAGCCACAGAG 754
Db 181 CCGTCAGGCGCCTGCTTCCAAAGTTTCATAGAGGCTCAGGATGAAGAGCCACAGAG 240

Qy 755 GGGACACGGCCACGCTGTGTGTGAGCTGAGCAGAGCGGCCACCGGTGGAGTGGAGGAAG 814
Db 241 GGGCCACAGCCACTACTGCGCTGTGAGCTGAGCAGAGCGGCCCGCTGGAGTGGAGGAAG 300

Qy 815 GGCATGAGACCTCAGAGATGGGACACACAGCCTCAGGCGAGCGGTCAGGTTG 874
Db 301 GGGCCGAGACCTCAGAGCGGGACAGAGTGGGGTGAAGCAGAGGAGCGGTGTG 360

Qy 875 AGCTGCAAGTCCGTGGCTGCTGTGGTGGATGCCGGGGAGTACTCGTGTGTGTCGGGC 934

Db 361 AGCTGGAGATCGTGGGCTGACCGTGGAGGATGCTGGGAGTACTCGTGCATGTGGGGC 420
Qy 935 AGGAGAGACCTCAGCCACACTCACTGTGAGGGCCCTCCTGCCAGATTCTAGAGATG 994
Db 421 AGGAGAGACCTCAGCCACAGTGGCGGTGAGGGCCCTGCTGCCAAGTTCTCTGAGGGTC 480

Qy 995 TGA 997
Db 481 TCA 483

RESULT 4
LOCUS BQ446463 723 bp mRNA linear EST 29-MAY-2002
DEFINITION UI-H-EUI-azz-h-22-0-UI.s1 NCI CGAP Ctl1 Homo sapiens cDNA clone
UI-H-EUI-azz-h-22-0-UI 3', mRNA sequence.
ACCESSION BQ446463
VERSION BQ446463
KEYWORDS BQ446463.1 GI:21249575
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 723)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jose Mercuende
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Clone distribution information can be obtained from Dr. M. Bento Soares, bento-soares@uiowa.edu
Seq primer: M13 FORWARD
POLYA=yes.
Location/Qualifiers
1. 723
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-H-EUI-azz-h-22-0-UI"
/tissue_type="Osteoarthritic Cartilage"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NCI CGAP Ctl1"
/note="Organ: Knee; Vector: p773-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP Ctl1 is a normalized cDNA library containing the following tissue(s): Osteoarthritic Cartilage The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dr primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p773-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is TGATCAGCT.
TAG TISSUE=osteoarthritis cartilage
TAG_LIB=UI-H-EUI
TAG_SEQ=TGATCAGCT"

ORIGIN

Query Match 30.2%; Score 302; DB 5; Length 723;
Best Local Similarity 68.0%; Pred. No. 1.1e-59;
Matches 449; Conservative 0; Mismatches 207; Indels 4; Gaps 2;

Qy 310 GACAGATACGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 369

Brunstein,A., deOliveira,P.S., Bucher,P., Brantani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 10737800
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-HN0020-221100-451-a02&t3=2000-11-22&t4=1)
 221100-451-a02&t3=2000-11-22&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 19
 High quality sequence stop: 437.
 Location/Qualifiers

FEATURES

1. 437
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HN0020"

/note="Organ: head normal; Vector: puc18; Site 1: Sma1;
 Site 2: Sma1; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 26.2%; Score 262.6; DB 2; Length 437;
 Best Local Similarity 80.4%; Pred. No. 1.5e-50;
 Matches 332; Conservative 0; Mismatches 79; Indels 2; Gaps 2;
 Oy 149 GGACTCAGCCACGCTCACCATCAGGCTCTGCCAGGTTTCATAGAGATGTGAAA 208
 Db 425 GGTCTCAGCTACACTCACTGTGAGGCGCTGCCAGACTCATAGAGATGTGAAA 366
 Oy 209 ACCAGGAGCCAGAGAGGGGCCACGGCTGTGCTGAGTGTGAGTGAACAGTGCAGCC 268
 Db 365 ACCAGAGGCCACAGAGGGGCTCAGTCACTTGCATGTAGCTGAGAAAGCGGCC 306
 Oy 269 CTGTGGAGTGGAGAAAGGGGTCTGAGACCTTCAGAGATGGGACAGATACAGCCTGAGGC 328
 Db 305 CCGTGGAGTGGAGAAAGGGGCCAACCCCTCAAGATGGGACAGGTACAGCCTGAAGC 246
 Oy 329 AGGACGGACTAATATGAGTGCAGATTCGTGGCTGGCCATGCGACACACTGGGAGT 388
 Db 245 AGGATGGACAGATTGTGAGTGCAGATTCTGGCTGTGATGAGTGTGAGAGTCTGAGAGAT 186
 Oy 389 ACTGTGGCTGTGGGCGAGAGAGGACCTCGGCTATGCTCACCCTGAGGCTCTACCCA 448
 Db 185 ACTGTGCATATGTAGCAGAGAGAGGACCTCGGCCACGCCCACTGTCTAGGGC-CTGCCGG 127
 Oy 449 TCAAGTTTCAAGAGGTTCTGAGGAACGAAGAGGCCACAGAAAGGGGCAACAGCCGTGTGTC 508
 Db 126 CCAGATTATAGAGATGTGAGAAATCACAGAGGCCACAGAGAGGGGCCACACAGCTGTGTGTC 67
 Oy 509 GGTGTAGCTGAGCAAGATGCCCCCTGGAGTGTGGTGAAGGGGCGATGAGACC 561
 Db 66 AGTGTAGCTGAGCACG-GCGCCCCCTGGAGTGTGGGAGGGGTCTGAACC 15

RESULT 7

BP826489

LOCUS CM4-HN0020-221100-452-b09 HN0020 Homo sapiens cDNA, mRNA linear EST 13-JAN-2001
 DEFINITION BP826489
 ACCESSION BP826489
 VERSION BP826489.1 GI:12169746
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE

1 (bases 1 to 369)
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bata,G.S., Simpson,D.H.,
 Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
 O'Hare,M.J., Soares,F., Brantani,R.R., Reis,L.F., de Souza,S.J. and
 Simpson,A.J.
 Shotgun sequencing of the human transcriptome with ORF expressed
 sequence tags
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
 20202663
 10737800
 Contact: Simpson A.J.G.
 Laboratory of Cancer Genetics
 Ludwig Institute for Cancer Research
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
 Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpson@ludwig.org.br
 This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
 (http://www.ludwig.org.br/scripts/gethtml2.pl?tl=CM4&t2=CM4-HN0020-221100-452-b09&t3=2000-11-22&t4=1)
 221100-452-b09&t3=2000-11-22&t4=1)
 Seq primer: puc 18 forward
 High quality sequence start: 37
 High quality sequence stop: 369.
 Location/Qualifiers

TITLE

JOURNAL
 MEDLINE
 PUBMED
 COMMENT

FEATURES

1. 369
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /dev_stage="Adult"
 /clone_lib="HN0020"

/note="Organ: head normal; Vector: puc18; Site 1: Sma1;
 Site 2: Sma1; A mini-library was made by cloning products
 derived from ORESTES PCR (U.S. Letters Patent application
 No. 196,716 - Ludwig Institute for Cancer Research)
 profiles into the pUC 18 vector. Reverse transcription of
 tissue mRNA and cDNA amplification were performed under
 low stringency conditions."

ORIGIN

Query Match 21.8%; Score 218.2; DB 2; Length 369;
 Best Local Similarity 81.4%; Pred. No. 3.3e-40;
 Matches 253; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
 Oy 2 CCCCTGTGGAGTGGAGGAAGGGGCCGAGAACCTCAGAGATGGGACAGATACATCTGA 61
 Db 52 CCCCGTTGGAGTGGATGAAGGGGCCCAACACCTCAAGATGGACAGGTACAGCCTGA 111
 Oy 62 GGACAGAGGGGACAGAGTGTGAGCTGCAGATCTGTGGCTGGCCATGGCGACCGCGGG 121
 Db 112 AGCAGGATGGGACAGATTGTGAGCTGCAGATTCTGTGGCTGTGATAGCAGATGCTGAG 171
 Oy 122 AGTACTTGTGTGTGTGGGCGAGGAGGACCTCAGCCAGCTCACCACATCAGGCTCTGC 181
 Db 172 AATACTGTGCATATGTGAGCAGGAGGAGGACCTCGGCCACCTCACTGTCTGGCCCTGC 231
 Oy 182 CTGCGAGGTTTCATAGAGATGTGAAAAACAGGAGGGCCAGAGAGGGGCCACGGCTGTGC 241
 Db 232 CGGCCAGATTTCATAGATATGTGAGAAATCACAGAGGCCACAGAGGGGCCACAGCTGTGC 291


```

QY 242 TGCAGTGTGAGTGAACAGTGCAGCCCCCTGTGTGAGTGGAGAAAGGGGTCTGAGACCCCTCA 301
Db |||||
292 TGCAGTGTGAGTGTGAGTGAACAGGCGGCCCCCTGTGTGAGTGGCGGAAGGGGTCTGAGACCCCTCA 351
QY 302 GAGATGGGAC 312
Db |||||
352 GAGCTGAGGAC 362

RESULT 8
CV315157 372 bp mRNA linear EST 24-SEP-2004
LOCUS CH1-HT0877-210900-418-b07 HT0877 Homo sapiens cDNA, mRNA sequence.
ACCESSION CV315157
VERSION CV315157.1 GI:52638371
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 372)
AUTHORS Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Balda,G.S., Simpson,D.H.,
Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,
O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the PAPESP/LICR Human Cancer Genome
Project. http://www.ludwig.org.br.
FEATURES
source
1..372
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0877"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI;
Site 2: SmaI; A mini-library was made by cloning products
derived from ORSTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
ORIGIN
Query Match 21.6%; Score 216; DB 7; Length 372;
Best Local Similarity 84.7%; Pred. No. 1.1e-39;
Matches 254; Conservative 0; Mismatches 45; Indels 1; Gaps 1;

QY 703 GCCATGCTTCCCAAGTTTCATAG-AGGGTCTGAGGAATGAAGAGGCCACAGAGGGGACAC 761
Db 65 GCGCTGCGGCCGATTTCATAGACGATGTGAGAAATCAAGAGGCCACAGAGGGGCGCAC 124

QY 762 GGCCACGCTGTGTGTGAGTGTGAGCAAGCGCGCACCGGTGAGTGGAGGAAGGGGATGA 821
Db 125 AGCTGTGCTGCAGTGTGAGCTGACAGAGCGGCCCCCGTGGAGTGGCGGAAGGGGTCTGA 184

QY 822 GACCTCAGAGATGGGACACAGACACGCTGAGCAGGACCGGTCACAGGTGTGAGCTGCA 881
Db 185 GACCTCAGAGATGGGACACAGATATCGCTGAGCAGGACCGGACGAGGTGTGAGCTGCA 244

```

```

QY 882 GATCGTGGCTGGCTGTGGTGGATGCCGGGAGTACTCTGTGTGTGTCGGGCGAGGAG 941
Db |||||
245 GATTCTGGCTGGCTGTGGAGGACACTGGAGAGTATTTGTGTGTGTCGGGCGAGGAG 304
QY 942 GACCTCAGCCACACTCACTGTTCAGGGCCCTGCCCTGCCAGATTTCATAGAAGATGTGAAAAA 1001
Db |||||
305 AACCTCAGCTACACTCACTGTTCAGGGCCCTGCCAGCCAGATTTCATAGACACATGACAAA 364

RESULT 9
CV316029 570 bp DNA linear GSS 30-SEP-2003
LOCUS tigr-gss-dog-17000332435861 Dog Library Canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE836029
VERSION CE836029.1 GI:37177049
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 570)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
FEATURES
source
1..570
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"
ORIGIN
Query Match 19.6%; Score 196.2; DB 9; Length 570;
Best Local Similarity 83.8%; Pred. No. 4.7e-35;
Matches 222; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 703 GCATGCGCTTCCCAAGTTTCATAGAGGTCTGAGGAATGAAGAGGCCACAGAGGGGACAG 762
Db 376 GCGCTGCGGCCCAAGTTTCACCAAGGCCCTGAGGAAGAGGAGGCCACCGAAGGGGCCAG 317

QY 763 GCCACGCTGTGTGTGAGCTGAGCAAGCGCGCACCGGTGGAGTGGAGGAAGGGGATGAG 822
Db 316 GCCACGCTGTGTGAGCTGAGCAAGCGCGGCCCGCTGGAGTGGAGGAAGGGGCTGAG 257

QY 823 ACCCTCAGAGATGGGGAACAGACACAGCTTGAAGCGGAGGACGGGTCCAGGTGTGAGCTGCAG 882
Db 256 GCGCTCAGAGCGCGGGAACAGAGGCTTGAAGCGGAGGACCGGGCCGTGTGCGAGTGCAG 197

QY 883 ATCCGTGGCGCTGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 942
Db 196 ATCCGGGGCGCTGTGCGTGTGAGGAGTCTCCGGGGAGTACTCGTGTGTGCGGGCAGGAGG 137

QY 943 ACCTCAGCCACACTCACTGTTCAGGG 967
Db 136 ACGGGCGCTGTGCTCACCATCAGGG 112

```


RESULT 10
BF853346/c
LOCUS
DEFINITION MR2-EN0091-191200-004-f08 EN0091 Homo sapiens CDNA, mRNA sequence. EST 16-JAN-2001
ACCESSION BF853346
VERSION BF853346.1 GI:12241090
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 359)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matekuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR2&t2=MR2-EN0091-191200-004-f08&t3=2000-12-19&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 19
High quality sequence stop: 358.
Location/Qualifiers
FEATURES
source
1. 359
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0091"
/note="Organ: lung normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 19.4%; Score 194.4; DB 2; Length 359;
Best Local Similarity 82.5%; Pred. No. 1.2e-34;
Matches 222; Conservative 0; Mismatches 47; Indels 0; Gaps 0;
QY 199 GATGTGAACACAGGAGCCAGAGAGGGCCACGGCTGTGCTGCAGTGTGAGCTGAC 258
DB 359 GATGTGAACATCAGAGGCCACAGAGAGGGCCACGGCTGTGCTGCAGTGTGAGCTGAGC 300
QY 259 AGTCGAGCCCTGTGTGAGTGGAGAAAGGGGTCTCGAGACCTTCAGAGATGGGGACAGATAC 318
DB 299 AAGNCGGCCCCGTGTGAGTGGCGAGGGGTCTGAGACCTTCAGAGATGGGGACAGATAT 240
QY 319 AGCCTGAGGACAGGACGGGACTTAATGTGAGCTGCAGATTCGTGGCTGGCCATGGCAGAC 378
DB 239 AGCCTGAGGACAGGACGGGACGAGGTGTGAGCTGCAGATTCGTGGCTGGCTGTGGAGGAC 180
QY 379 ACTGGGAGTACTCGTGGCTGGGGCAGAGAGGACCTCGGTATGCTCACCGTCAGG 438

Db 179 ACTGGAGAGTATTGTGTGTGCGGGCAGAGAGAACCTCAGCTACACTACTGTCTCAGG 120
QY 439 GCTTACCCATCAAGTTCACAGAGGGTCT 467
Db 119 GGTAAGATCCTGTGTGTGTCACATGGACT 91
RESULT 11
LOCUS
DEFINITION MR2-EN0387-290900-001-g11 BN0387 Homo sapiens CDNA, mRNA sequence. EST 24-SEP-2004
ACCESSION CV346932
VERSION CV346932.1 GI:52670146
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 515)
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F., Goldman, G.H., Carvalho, A.F., Matekuma, A., Baia, G.S., Simpson, D.H., Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. http://www.ludwig.org.br.
Location/Qualifiers
FEATURES
source
1. 515
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="BN0387"
/note="Organ: breast normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."
ORIGIN
Query Match 18.6%; Score 186; DB 7; Length 515;
Best Local Similarity 78.7%; Pred. No. 1.1e-32;
Matches 222; Conservative 0; Mismatches 60; Indels 0; Gaps 0;
QY 586 CCATGCTCACCGTCAGGGCCATGCTTCCAGTTTCATAGAGGGTCTGAGGATGAGAGG 745
DB 172 CCATGTGACCCCTTCCTAGCCCTGCCTGCACGATTCATCAAGATCTGAAGACCAAGAGG 231
QY 746 CCACAGAAGGGGACACGGCCACGCTGTGTGTGAGCTGAGCAAGCGGCACCGTGGAGT 805
DB 232 CTTCAAGAGGGGCCACAGCTACACTGCAGTGTGAGCTGAGCAAGGTGGCCCTGTGGAT 291
QY 806 GGAGGAAGGGGCATGAGACCTTCAGAGATGGGGAACAGACAGCTGAGGACGAGCGGT 865
DB 292 GGAAGAAGGGTCTCTGAGACCTCTAGAGATGGGGCAGATACAGCTCTGAAGCAGGATGGA 351
QY 866 CCAGGTGTGACTGCAGATCCGTGGCTGGTGTGGATGCCGGGGAGTACTCTGTGTG 925


```

Db      352  CGAGGCGTGCAGATCATGACCTGTCGTGGCGGATGCTGGGAATCTCATGCA 411
QY      926  TGTGCGGCGAGGAGACCTCAGCCACACTACTGTTCAGGG 967
Db      412  TGTGTGCAAGAGAGAGACCTCGGCCACGCTCAGGTTCAGGG 453

RESULT 12
LOCUS   BF398067/c
DEFINITION UI-R-BS2-beg-h-06-0-UI.s1 UI-R-BS2 Rattus norvegicus cDNA clone
ACCESSION BF398067
VERSION   BF398067.1 GI:11383052
KEYWORDS  EST.
SOURCE    Rattus norvegicus (Norway rat)
ORGANISM  Rattus norvegicus
REFERENCE 1 (bases 1 to 601)
AUTHORS   Ronaldo,M.F., Lennon,G. and Soares,M.B.
TITLE     Normalization and subtraction: two approaches to facilitate gene
          discovery
JOURNAL   Genome Res. 6 (9), 791-806 (1996)
MEDLINE   97044477
PUBMED    8889548
COMMENT   Contact: Soares, MB
          Coordinated Laboratory for Computational Genomics
          University of Iowa
          375 Newton Road , 4156 MEBRF, Iowa City, IA 52242, USA
          Tel: 319 335 8250
          Fax: 319 335 9565
          Email: bento-soares@uiowa.edu
          The sequence contained an oligo-dT track that was present in the
          oligonucleotide that was used to prime the synthesis of first
          strand cDNA and therefore this may represent a bonafide poly A
          tail. The sequence tag present in the cDNA between the NotI site
          and the oligo-dT track served to verify it as a clone from the
          normalized embryo at 13 dpc library cDNA Library preparation: M.B.
          Soares Lab Clon distribution: clones will be available through
          Research Genetics (www.resgen.com)
          Seq primer: M13 Forward
          POLYA=Yes.

FEATURES
          source
          1..601
             Location/Qualifiers
             /organism="Rattus norvegicus"
             /mol_type="mRNA"
             /strain="Sprague-Dawley"
             /db_xref="taxon:10116"
             /clone="UI-R-BS2-beg-h-06-0-UI"
             /dev_stages="embryonic 13 dpc"
             /lab_host="DH10B (Life Technologies)"
             /clone_lib="UI-R-BS2"
             /note="Vector: pRT3D-Pac (Pharmacia) with a modified
             polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BS2
             library is a subtracted library derived from 13 dpc whole
             embryo tissue. For a detailed description of the library
             from which this clone was derived, please visit our web
             site at ratest.eng.uiowa.edu. The subtraction has been
             previously described in (Bonaldo, Lennon and Soares,
             Genome Research 6:791-806, 1996)
             TAG_TISSUE=embryo at 13 dpc
             TAG_L1B=UI-R-BS2
             TAG_SEQ=AATCC"

ORIGIN
Query Match 17.5%; Score 175.2; DB 2; Length 601;
Best Local Similarity 71.3%; Pred. No. 3.7e-30;
Matches 231; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY      644  CTGGGGAGTACTGTGATGTGCGGAAGAGAGACCTCAGCCATCTCAGGTTCAGGG 703

```

```

Db      517  CTAGGGGCGTCCATGTTCTCAGAGGGTGTGATTGTGTCTGTGTCTTTCTGACCTCCCCAG 458
QY      704  CCATGCCCTTCCAAGTTCATAGAGGGTCTGAGAGATGAAGAGGCCACAGAGAGGGGACACGG 763
Db      457  CCCITTCGGGCTAAGTTCATAGAAAGTCTGAAGATGAAGGGGCCACAGAGGAACCAACAG 398
QY      764  CCAGCTGTGTGTGAGCTGAGCAAGGCGGCACCGGTGGAGTGGAGGAAGGGGCATGAGA 823
Db      397  CCACGCTGAGCTGCAAACTGAGCAAGGCGGCTCCGGTGAAGCTGGAAGAGAGGAACAAGA 338
QY      824  CCCTCAGAGATGGGGACAGACACAGCCCTGAGGAGGAGCGGTTCAGGTGTGAGTTCAGA 883
Db      337  CCTTGCAGAGCGGAGACAATATGCGCTGAGGCGAGGACGGAGCTGTGTGTGAGCTGAGA 278
QY      884  TCCGTGGCCTGTGTGTGTGGATGCCGGGGAGTACTCTGTGTGTGCGGGCAGGAGAGGA 943
Db      277  TCCGTGGCCTGACACAGCCGATGCTGGGAGTACTCATGTGTGTGTGGGACAGGAGAGA 218
QY      944  CCTCAGCCACTCCTCCTGTTCAGGG 967
Db      217  CATCAGCTGCTGACTGTCAAGG 194

RESULT 13
LOCUS   BQ361458/c
DEFINITION PMO-OT0232-240501-005-a12 OT0232 Homo sapiens cDNA, mRNA sequence.
ACCESSION BQ361458
VERSION   BQ361458.1 GI:21031179
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
REFERENCE 1 (bases 1 to 324)
AUTHORS   Nagai,M,A., da Silva,W. Jr., Zago,M,A., Bordin,S., Costa,F.F., Briones,M.R.,
          Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
          Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
          O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
          Simpson,A.J.
          Shotgun sequencing of the human transcriptome with ORF expressed
          sequence tags
          Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
          20202663
          10737800
          Contact: Simpson A.J.G.
          Laboratory of Cancer Genetics
          Ludwig Institute for Cancer Research
          Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
          Brazil
          Tel: +55-11-2704922
          Fax: +55-11-2707001
          Email: asimpson@ludwig.org.br
          This sequence was derived from the FAPESP/LICR Human Cancer Genome
          Project. This entry can be seen in the following URL
          (http://www.ludwig.org.br/scripts/gethtml2.pl?ci=PMO&t2=PMO-OT0232-
          240501-005-a12&t3=2001-05-24&t4=1)
          Seq primer: puc 18 forward
          High quality sequence start: 21
          High quality sequence stop: 324.
          Location/Qualifiers
          1..324
             /organism="Homo sapiens"
             /mol_type="mRNA"
             /db_xref="taxon:9606"
             /dev_stage="Adult"
             /clone_lib="OT0232"
             /note="Organ: ovary; Vector: puc18; Site 1: SmaI; Site 2:
             SmaI; A mini-library was made by cloning products derived
             from: ORESTES PCR (U.S. Letters Patent application No.
             196,716 - Ludwig Institute for Cancer Research) profiles

```


into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

ORIGIN
Query Match          17.3%; Score 173.2; DB 5; Length 324;
Best Local Similarity 89.0%; Pred. No. 9.9e-30;
Matches 187; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

Qy 757 GACAGCGCCAGCTGTGTGAGCTGAGCAAGCGGCCACCGTGGAGTGGAGGAAGGGG 816
Db 324 GCCAGCGCCAGCTGTGTGAGCTGAGCAAGCGGCCACCGTGGAGTGGAGGAAGGGG 265

Qy 817 CATGAGACCTTCAGAGATGGGACACAGACACAGCTGAGGACGAGCGGTCCAGGTGTGAG 876
Db 264 TCCGAGACCTTCAGAGATGGGACACAGATCTGTCTGAGGACGAGCGGTCCAGGTGTGAG 205

Qy 877 CTGACAGATCCGTGGCTGTGTGTGATGCGCGGAGTACTCTGTGTGTGTGCGGGCAG 936
Db 204 CTGACAGATCCGTGGCTGTGTGTGATGCGCGGAGTACTCTGTGTGTGTGCGGGCAG 145

Qy 937 GAGAGGACTCAGCCACACTACTCTCAGG 966
Db 144 GAGAGGACTCAGCCCTACTCACCATCAGG 115

```

```

RESULT 14
BH025901/c
LOCUS          BH025901          690 bp      DNA      linear      GSS      17-JUL-2001
DEFINITION    RPCI-24-259H14-TV RPCI-24 Mus musculus genomic clone
ACCESSION     BH025901
VERSION       BH025901.1      GI:14789365
KEYWORDS      GSS.
SOURCE        Mus musculus (house mouse)
ORGANISM      Mus musculus

```

```

REFERENCE
1 (bases 1 to 690)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Zhao, S., Nierman, W., Malek, J., Shatsman, S., Akinret, B., Levins, M.,
Tsagay, G., Geer, K., Krol, M., Shvartsbeyn, A., Gebregorgis, E.,
Russell, D., de Jong, P. and Fraser, C.M.
Mouse BAC End Sequences from Library RPCI-24
Unpublished (1999)
Other GSSs: RPCI-24-259H14.TJ
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org

```

```

Clones are derived from the mouse BAC library RPCI-24. For BAC
library availability, please contact Pieter de Jong
(pdejong@mail.cho.org). Clones may be purchased from BACPAC
Resources (http://www.choi.org/bacpac/orderingframe.html). BAC end
page: http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 259 row: H column: 14
Seq primer: T7
Class: BAC ends.

```

```

FEATURES
source
1..690
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-259H14"
/sex="Male"
/cell_type="Spleen/Brain"
/clone_lib="RPCI-24"
/note="Vector: pTARBAC1; Site 1: BamHI; Site 2: BamHI;
RPCI-24 Mouse BAC Library produced by Pieter de Jong. The
library was cloned in the pTARBAC1 cloning vector at the

```

BamHI sites using MboI partially digested male C57BL/6J DNA."

```

ORIGIN
Query Match          17.3%; Score 173.2; DB 8; Length 690;
Best Local Similarity 82.7%; Pred. No. 1.1e-29;
Matches 210; Conservative 0; Mismatches 43; Indels 1; Gaps 1;

Qy 703 GCCATGCTTCAAGTTTCATAGAGGCTCTGAGGAATGAAGAGGCCACAGAGGGGACAG 762
Db 253 GCCCTGCCCTCCAGATTCATAGAGATTTGAGAAGTCAAGAGGCCACAGAGGCCACCATG 194

Qy 763 GCCAGCTGTGTGTGAGCTGAGCAAGCGGCCACCGTGGAGTGGAGNAGGGGCTAG 822
Db 193 GCCACTCTGAGGTGCCAGATGAGCAAGACTGCCCTGTGGAGTGAAGAAGGGGTCTGAG 134

Qy 823 ACCCTCAGAGATGGGGACAGACACAGCTGAGGACGAGCGGTCCAGGTGTGAGCTGAG 882
Db 133 ACCCTGAGAGATGGGGCAGATACAGCTGAGGACGAGTGGCTGTGTGTGAGCTGAG 74

Qy 883 ATCGTGCGCTGGCTGTGTGTGATGCCGGGAGTACTCTGTGTGTGTGCGGCGAGAGAGG 942
Db 73 ATCTGTGACCTGGCTGTGCAAGAGCGCTGGGAGTACTCATGTGTGTG-GGGCAGGAGAAG 15

Qy 943 ACCTCAGCCACACT 956
Db 14 ACCTCAGCCACACT 1

```

```

RESULT 15
BG987697/c
LOCUS          BG987697          204 bp      mRNA      linear      EST      13-JUN-2001
DEFINITION    PM3-HT1165-110101-001-e07 HT1165 Homo sapiens cDNA, mRNA sequence.
ACCESSION     BG987697
VERSION       BG987697.1      GI:14391767
KEYWORDS      EST.
SOURCE        Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

REFERENCE
1 (bases 1 to 204)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Dias Neto, E., Garcia Correa, R., Vertovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, P.F.,
Goldman, G.H., Carvalho, A.P., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.

```

```

Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br

```

```

This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=PM3&t2=PM3-HT1165-
110101-001-e07&t3=2001-01-11&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 35
High quality sequence stop: 204.
Location/Qualifiers
1..204
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"

```

```

FEATURES
source

```


/clone_lib="HT1165"
/note="Organ: head neck; Vector: puc18; Site 1: Sma1;
Site 2: Sma1; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."

ORIGIN

```
Query Match      14.7%; Score 147; DB 4; Length 204;  
Best Local Similarity 82.8%; Pred. No. 1.2e-23;  
Matches 168; Conservative 0; Mismatches 35; Indels 0; Gaps 0;  
  
Qy 57 CCTGAGGCGAGGAGGACCAAGGTGTGAGCTGCAGATCTGTGGCTGGCCATGGCGGACGC 116  
    |||||  
Db 204 CCTGAGGCGAGGAGGAGGAGGTGTGAGCTGCAGATCTGTGGCTGGCTGTGGAGGACAC 145  
    |||||  
  
Qy 117 CGGGGAGTACTTGTGTGTGTGGGGCAGGAGGAGGACCTCAGCCACGCTCACCATCAGGGC 176  
    |||||  
Db 144 TGGAGAGTATTGTGTGTGTGTGGGGCAGGAGGAGGACCTCAGCTACACTCACTTTCAGGGC 85  
    |||||  
  
Qy 177 TCTGCCCTGCCAGGTTCATAGAGATGTGAAACACAGGAGGCGCCAGAGAGGGGCCACGGC 236  
    |||||  
Db 84 CCTGCCAGCCAGATTTCATACCCCATGACCAACAGGAAGCCAGAGAGGGGCCACGGC 25  
    |||||  
  
Qy 237 TGTGCTGCAGTGTGAGCTGAACA 259  
    |||||  
Db 24 CACACTGCACTGTGAACCTGAGCA 2
```

Search completed: March 21, 2005, 14:13:38
Job time : 3668.47 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:24:10 ; Search time 4515.62 Seconds
(without alignments)
10741.310 Million cell updates/sec

Title: US-10-077-130-4_COPY_22500_23500
Perfect score: 1001
Sequence: 1 ggcagtcagtgacactggc.....ctgtcgaagtcgtggacct 1001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_hcg.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pi.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_av.*
- 13: gb_un.*
- 14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	3225	6	AX642969 Sequence
2	1001	100.0	4041	9	AB046859 Homo sapi
3	1001	100.0	4175	6	AX039409 Sequence
4	1001	100.0	4936	6	AX207401 Sequence
5	1001	100.0	5454	6	AX504255 Sequence
6	999.4	99.8	5007	6	AX039410 Sequence
7	999.4	99.8	5207	6	AR255867 Sequence
8	999.4	99.8	5207	6	AR442826 Sequence
9	999.4	99.8	5207	6	AR442827 Sequence
10	999.4	99.8	5207	6	AX430856 Sequence
11	999.4	99.8	5207	6	AX430858 Sequence
12	999.4	99.8	7928	6	AX039412 Sequence
13	754.6	75.4	18524	6	CQ730656 Sequence
14	738.4	73.8	4740	10	AY603754 Mus muscu
15	406.6	40.6	1183	10	BC046431 Mus muscu
16	243	24.3	62164	9	AL670729 Human DNA
c 17	243	24.3	174612	2	AC023889 Homo sapi
18	184.2	18.4	5923	5	AY603753 Danio rer
c 19	181	18.1	164766	2	AC026657 Homo sapi

c 20	180.6	18.0	103129	10	AL662809	AL662809 Mouse DNA
c 21	180.6	18.0	258622	2	AC020877	AC020877 Mus muscu
c 22	174.2	17.4	260998	2	AC099089	AC099089 Rattus no
23	164.8	16.5	182	6	AX913866	AX913866 Sequence
24	164.8	16.5	182	6	BD049399	BD049399 Sequence
25	156.4	15.6	4357	10	AK122488	AK122488 Mus muscu
26	150	15.0	6726	9	AB037718	AB037718 Homo sapi
27	150	15.0	7710	6	AX056397	AX056397 Sequence
28	150	15.0	9282	9	AY603755	AY603755 Homo sapi
29	148.4	14.8	7789	6	AX746187	AX746187 Sequence
30	143.8	14.4	10801	10	AF215896	AF215896 Mus muscu
31	142.2	14.2	6847	6	C0714785	C0714785 Sequence
32	115.4	11.5	250087	2	AC096931	AC096931 Rattus no
33	108	10.8	98613	2	AC139630	AC139630 Takifugu
34	107	10.7	176384	10	AC114651	AC114651 Mus muscu
c 35	107	10.7	184776	2	AC115011	AC115011 Bos tauru
c 36	106.8	10.7	185902	2	AC150987	AC150987 Bos tauru
c 37	106.8	10.7	192547	2	AC150990	AC150990 Bos tauru
38	105.6	10.5	183625	9	AC053503	AC053503 Homo sapi
39	100.6	10.0	178817	2	CR394562	CR394562 Danio rer
40	100.6	10.0	184176	2	AC149120	AC149120 Papio anu
41	100.6	10.0	193064	2	AC149860	AC149860 Papio anu
c 42	100.4	10.0	216648	10	AC112361	AC112361 Rattus no
43	100.4	10.0	251133	2	AC121633	AC121633 Rattus no
44	95.6	9.6	159699	2	AC151375	AC151375 Callithri
45	94.6	9.5	9530	5	AY578914	AY578914 Danio rer

ALIGNMENTS

RESULT 1
AX642969
LOCUS AX642969 3225 bp DNA linear PAT 24-FEB-2003
DEFINITION Sequence 46 from Patent WO01096547.
ACCESSION AX642969
VERSION AX642969.1 GI:28550118
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Yue, H., Lal, P., Bandman, O., Borowsky, M., Au-Young, J., Lu, Y.,
Gandhi, A.R., Tribouley, C.M., Wallia, N., Yao, M.G., Lu, D.A.,
Greenwald, S.R., Ramkumar, J., Griffin, J.A., Kearney, L., Burford, N.,
Nguyen, D.B., Tang, Y.T., Baughn, M.R., He, A., Thornton, M.,
Hafalia, A., Patterson, C., Gururajan, R., Lo, T.P., Khan, F.,
Recipon, S.A., Azimzai, Y., Policky, J.L., Ding, L., Grether, M.,
Elliott, V.S., Thangavelu, K., Batra, S. and Ison, C.H.
HUMAN KINASES
TITLE Human kinases
JOURNAL Patent: WO 01096547-A 46 20-DEC-2001;
Incyte Genomics, Inc. (US)
FEATURES
source Location/Qualifiers
1..3225
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 3016969CBI"

Query Match 100.0%; Score 1001; DB 6; Length 3225;
Best Local Similarity 100.0%; Pred. No. 5.3e-156;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GGCCAGTCAGTGACACTGGCCTGCCAGTGTCAGCCAGCCAGCTGCCAGCCACCTGGG	60
Db	1663	GGCCAGTCAGTGACACTGGCCTGCCAGTGTCAGCCAGCCAGCTGCCAGCCACCTGGG	1722
Qy	61	AGCAAAAGCAGGAGCCCTCGGAGAGCAGCAGCGCTGTCCTCATCTCTGCCACCTCAAG	120
Db	1723	AGCAAAAGCAGGAGCCCTCGGAGAGCAGCAGCGCTGTCCTCATCTCTGCCACCTCAAG	1782

QY	481	TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGCGCCAGCCACCTGGCCCTGAGGAG	540
Db	3058	TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGCGCCAGCCACCTGGCCCTGAGGAG	3117
QY	541	GAGAGCAGGGCGGTGAGCCAAACCCCTGCCAGCAAAAAGACCTTCGATTCAGACA	600
Db	3118	GAGAGCAGGGCGGTGAGCCAAACCCCTGCCAGCAAAAAGACCTTCGATTCAGACA	3177
QY	601	CAGATCCAGAGGGCGGTGAGCCAAACCCCTGCCAGCAAAAAGACCTTCGATTCAGACA	660
Db	3178	CAGATCCAGAGGGCGGTGAGCCAAACCCCTGCCAGCAAAAAGACCTTCGATTCAGACA	3237
QY	661	GGCTGGCCGCAAGATCATCCCTTACCAACCCCAAGCAAGACAGCAGTGTGCGCGAA	720
Db	3238	GGCTGGCCGCAAGATCATCCCTTACCAACCCCAAGCAAGACAGCAGTGTGCGCGAA	3297
QY	721	TAGGAGCCCTCAGGGCCCTGGCCACCCGCACTCTGGCCGAGCAGTGTGCGCGAA	780
Db	3298	TAGGAGCCCTCAGGGCCCTGGCCACCCGCACTCTGGCCGAGCAGTGTGCGCGAA	3357
QY	781	AGCCCCCGGCACCTGGTGTCTCATCTTGGAGCTGTCTCTGGCCGAGCAGTGTCTCCCTGC	840
Db	3358	AGCCCCCGGCACCTGGTGTCTCATCTTGGAGCTGTCTCTGGCCGAGCAGTGTCTCCCTGC	3417
QY	841	CTGGCCGAGAGGCGCTCTCTACTCAGAACTCCGAGGTGAAGGACTACCTGTGGCAGATGTTG	900
Db	3418	CTGGCCGAGAGGCGCTCTCTACTCAGAACTCCGAGGTGAAGGACTACCTGTGGCAGATGTTG	3477
QY	901	AGTGCACCCAGTACTGAGCAACAGACATCTCTGACCTGGAGCTGAGGTCCGAGAAC	960
Db	3478	AGTGCACCCAGTACTGAGCAACAGACATCTCTGACCTGGAGCTGAGGTCCGAGAAC	3537
QY	961	ATGATCATCACCGAATAACAACCTGCTCAAGGTCTGGACCT	1001
Db	3538	ATGATCATCACCGAATAACAACCTGCTCAAGGTCTGGACCT	3578
RESULT 4			
AX207401			
LOCUS			
DEFINITION			
AX207401			
ACCESSION			
VERSION			
AX207401.1			
GI:15395213			
KEYWORDS			
Homo sapiens (human)			
SOURCE			
Homo sapiens			
ORGANISM			
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
REFERENCE			
1			
Plowman, G., Whyte, D., Manning, G., Sudarsanam, S., and Martinez, R.			
Human protein kinases and protein kinase-like enzymes			
Patent: WO 0155356-A 14 02-AUG-2001;			
Sugen, Inc. (US)			
FEATURES			
Location/Qualifiers			
1..4936			
/organism="Homo sapiens"			
/mol_type="unassigned DNA"			
/db_xref="taxon:9606"			
ORIGIN			
Query Match			
Best Local Similarity			
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY			
1			
GGCCAGTCAGTGACATGGCTGCGCAGTGTGAGCCAGCCAGCAGTGTGCGCCAGCCAGTGTG			
Db			
3379			
GGCCAGTCAGTGACATGGCTGCGCAGTGTGAGCCAGCCAGCAGTGTGCGCCAGCCAGTGTG			
QY			
61			
AGCAAGAGCAGGAGCCCTGGAGAGCAGCAGCGTGTCTCTATCTCTGCAACCTCAAG			
Db			
3439			
AGCAAGAGCAGGAGCCCTGGAGAGCAGCAGCGTGTCTCTATCTCTGCAACCTCAAG			

QY	121	AAC	TCCAGCTTCTGACCATCTGGTGGTGGCTGAGGACCTGGGTGTGTACACCTGC	180
Db	3499	AAC	TCCAGCTTCTGACCATCTGGTGGTGGCTGAGGACCTGGGTGTGTACACCTGC	3558
QY	181	AGCG	TGAGCAATGGCTGGGGACAGTGACCAACACGGGCGTCTCCGGAAGGCAGAGCGC	240
Db	3559	AGCG	TGAGCAATGGCTGGGGACAGTGACCAACACGGGCGTCTCCGGAAGGCAGAGCGC	3618
QY	241	CCCT	CATCTTTCCGATATGCCCGGATATCGGGAGGTGTACGCGATGGGTGTCTGTGTC	300
Db	3619	CCCT	CATCTTTCCGATATGCCCGGATATCGGGAGGTGTACGCGATGGGTGTCTGTGTC	3678
QY	301	TGGN	AGCCCGTGGAAATCTACGGCCCTGTGACCTACATTTGTCAGTGTGAGTGTGAGGC	360
Db	3679	TGGN	AGCCCGTGGAAATCTACGGCCCTGTGACCTACATTTGTCAGTGTGAGTGTGAGGC	3738
QY	361	GGCAG	CTGGACCACTGGGCTCCGACATCTTTGACTGTCTGCTACCTGACAGCAAGCTC	420
Db	3739	GGCAG	CTGGACCACTGGGCTCCGACATCTTTGACTGTCTGCTACCTGACAGCAAGCTC	3798
QY	421	TCCG	GGGGTGGCACTTACCTTCCGCAAGGATGTGTAGCAAGGAGGAGTGGGTCCC	480
Db	3799	TCCG	GGGGTGGCACTTACCTTCCGCAAGGATGTGTAGCAAGGAGGAGTGGGTCCC	3858
QY	481	TACAG	CAGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCACCTGGCTCTGAGGAG	540
Db	3859	TACAG	CAGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCACCTGGCTCTGAGGAG	3918
QY	541	GAGAG	CAGGGGGGTGAGCCCAACCCCTGCCAGCAAAAGACCTTTCGATTTCCAGACA	600
Db	3919	GAGAG	CAGGGGGGTGAGCCCAACCCCTGCCAGCAAAAGACCTTTCGATTTCCAGACA	3978
QY	601	CAGAT	CCAGAGGGGGCGCTTTCAGGTGGTGGGCAATGCTGGGAGAGGCGAGGGGCGG	660
Db	3979	CAGAT	CCAGAGGGGGCGCTTTCAGGTGGTGGGCAATGCTGGGAGAGGCGAGGGGCGG	4038
QY	661	GGCT	GGGCCGCCAAGATCATCTCCCTTACCAACCCCAAGGACAGCAGTGTCTGCGCGAA	720
Db	4039	GGCT	GGGCCGCCAAGATCATCTCCCTTACCAACCCCAAGGACAGCAGTGTCTGCGCGAA	4098
QY	721	TAGAG	GGCCCTCAAGGGCTTGGCCACCCGACCTGGCCAGTGTGACCGAGCTTACCTC	780
Db	4099	TAGAG	GGCCCTCAAGGGCTTGGCCACCCGACCTGGCCAGTGTGACCGAGCTTACCTC	4158
QY	781	AGCC	CGCGACCTGTGTCTCATCTTGGAGCTGTGCTCTGGGCGCCAGTGTCTCCCTGC	840
Db	4159	AGCC	CGCGACCTGTGTCTCATCTTGGAGCTGTGCTCTGGGCGCCAGTGTCTCCCTGC	4218
QY	841	CTGG	CGGAGAGGGCTCTCTACTCAGAACTCCGAGGTGAAGGACTTACCTGTGGCAGATGTTG	900
Db	4219	CTGG	CGGAGAGGGCTCTCTACTCAGAACTCCGAGGTGAAGGACTTACCTGTGGCAGATGTTG	4278
QY	901	AGTG	CCACCCAGTACTGTCGACAAACAGACATCTCTGCACTGGACCTGAGGTCCGAGAAC	960
Db	4279	AGTG	CCACCCAGTACTGTCGACAAACAGACATCTCTGCACTGGACCTGAGGTCCGAGAAC	4338
QY	961	ATGAT	CATCACCGAATAACAACCTGCTCAAGGTCTGGGACCT	1001
Db	4339	ATGAT	CATCACCGAATAACAACCTGCTCAAGGTCTGGGACCT	4379
RESULT 5				
AX504255				
LOCUS				
DEFINITION				
Sequence 44 from Patent WO0233099.				
AX504255				
ACCESSION				
VERSION				
AX504255.1				
GI:23386095				
KEYWORDS				
Homo sapiens (human)				
SOURCE				
Homo sapiens				
ORGANISM				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE				
1				

AUTHORS	Gururajan, R., Baughn, M.R., Walia, N.K., Elliott, V.S., Xu, Y., Arvizu, C., Yao, M.G., Ramkumar, J., Ding, L., Tang, Y.T., Hafalia, A.J., Nguyen, D.B., Gandhi, A.R., Lu, Y., Yue, H., Burford, N., Bandman, O., Tribouley, C.M., Lai, P.G., Recipon, S.A., Lu, D.A., Borowsky, M.L., Thornton, M., Swarnaker, A., Thangavelu, K., Khan, F.A. and Ison, C.H.																				
TITLE	Human kinases																				
JOURNAL	Patent: WO 0233099-A 44 25-APR-2002;																				
FEATURES	Incyte Genomics, Inc. (US)																				
source	Location/Qualifiers																				
	1. 5454																				
	/organism="Homo sapiens"																				
	/mol_type="unassigned DNA"																				
	/db_xref="taxon:9606"																				
	/notes="Incyte ID No: 7638121CB1"																				
ORIGIN																					
Query Match	100.0%; Score 1001; DB 6; Length 5454;																				
Best Local Similarity	100.0%; Pred. No. 4.7e-156;																				
Matches 1001; Conservative	0; Mismatches 0; Indels 0; Gaps 0;																				
Qy	1	GGCCAGT	CAGT	GAC	CTGG	CTGC	AGGT	GT	CAG	CC	CAG	CAG	CTG	CC	C	AGG	CC	AC	CTG	G	60
Db	3892	GGCCAGT	CAGT	GAC	CTGG	CTGC	AGGT	GT	CAG	CC	CAG	CAG	CTG	CC	C	AGG	CC	AC	CTG	G	3951
Qy	61	AGCAAG	ACGG	AG	CCCC	CTTG	AGAG	CAG	CAG	CG	CTG	CT	CA	T	CT	GC	CA	CC	CT	CA	120
Db	3952	AGCAAG	ACGG	AG	CCCC	CTTG	AGAG	CAG	CAG	CG	CTG	CT	CA	T	CT	GC	CA	CC	CT	CA	4011
Qy	121	AAC	TT	CAG	CTT	CTG	AC	CA	CT	CTG	TGG	TGG	CTG	AG	CA	CC	TGG	GT	GT	TAC	180
Db	4012	AAC	TT	CAG	CTT	CTG	AC	CA	CT	CTG	TGG	TGG	CTG	AG	CA	CC	TGG	GT	GT	TAC	4071
Qy	181	AG	CG	TG	AG	CA	AT	GG	CG	TG	GG	GA	CAG	TG	CA	CA	CC	AG	GG	CG	240
Db	4072	AG	CG	TG	AG	CA	AT	GG	CG	TG	GG	GA	CAG	TG	CA	CA	CC	AG	GG	CG	4131
Qy	241	CC	CT	CA	T	CT	TC	CG	CA	T	CC	CG	GA	T	AT	CG	GG	AG	GG	CT	300
Db	4132	CC	CT	CA	T	CT	TC	CG	CA	T	CC	CG	GA	T	AT	CG	GG	AG	GG	CT	4191
Qy	301	TGG	AG	CC	CT	GG	CA	CT	T	AG	CC	CT	GT	GA	CC	T	TG	CA	GT	GA	360
Db	4192	TGG	AG	CC	CT	GG	CA	CT	T	AG	CC	CT	GT	GA	CC	T	TG	CA	GT	GA	4251
Qy	361	GG	CA	GT	TG	GA	CC	CA	CA	CT	GG	CG	CT	GT	GA	CC	T	TG	CA	GT	420
Db	4252	GG	CA	GT	TG	GA	CC	CA	CA	CT	GG	CG	CT	GT	GA	CC	T	TG	CA	GT	4311
Qy	421	T	CC	CG	GG	TG	CA	CT	T	CC	GA	CG	CA	T	GT	CT	CA	GA	AG	CA	480
Db	4312	T	CC	CG	GG	TG	CA	CT	T	CC	GA	CG	CA	T	GT	CT	CA	GA	AG	CA	4371
Qy	481	T	AC	AG	CA	CC	CT	CG	GA	CA	AG	T	CT	CT	CG	GA	GG	CC	CA	CC	540
Db	4372	T	AC	AG	CA	CC	CT	CG	GA	CA	AG	T	CT	CT	CG	GA	GG	CC	CA	CC	4431
Qy	541	G	AG	AG	CC	GG	CG	GT	C	AG	CC	CA	CC	CT	TG	CC	AG	CA	CA	AG	600
Db	4432	G	AG	AG	CC	GG	CG	GT	C	AG	CC	CA	CC	CT	TG	CC	AG	CA	CA	AG	4491
Qy	601	C	AG	T	CC	AG	GG	CC	CG	CT	T	C	AG	CG	T	GG	CA	A	T	G	720
Db	4492	C	AG	T	CC	AG	GG	CC	CG	CT	T	C	AG	CG	T	GG	CA	A	T	G	4551
Qy	661	G	CG	CT	G	CG	CA	GA	T	AT	CC	CT	CA	CC	CA	AG	GA	CA	AG	CA	720
Db	4552	G	CG	CT	G	CG	CA	GA	T	AT	CC	CT	CA	CC	CA	AG	GA	CA	AG	CA	4611
Qy	721	T	AC	GA	GC	CT	CA	AG	GG	CC	TG	CG	CA	A	T	G	CT	GC	CA	GC	780
Db</																					

Query Match 99.8%; Score 999.4; DB 6; Length 5007;
Best Local Similarity 99.9%; Pred. No. 8.8e-156;
Matches 1000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCAGTCAGTGACACTGGCTCCAGGTGTACGCCAGCAGCTGCCAGGCCACTGG 60
Db |||||||
3451 GGCAGTCAGTGACACTGGCTCCAGGTGTACGCCAGCAGCTGCCAGGCCACTGG 3510
Qy 61 AGCAAGACGAGGCCCTCCAGGTGTACGCCAGCAGCTGCCAGGCCACTGG 120
Db |||||||
3511 AGCAAGACGAGGCCCTCCAGGTGTACGCCAGCAGCTGCCAGGCCACTGG 3570
Qy 121 AACTTCAGCTTCTGACCAATCTGGTGTGGTGTAGGACCTGGGTGTACACCTGC 180
Db |||||||
3571 AACTTCAGCTTCTGACCAATCTGGTGTGGTGTAGGACCTGGGTGTACACCTGC 3630
Qy 181 AGCGTCAGCAATGCGCTGGGACAGTGACCAACCGGGGTCTCTCGGAAGCAGAGGC 240
Db |||||||
3631 AGCGTCAGCAATGCGCTGGGACAGTGACCAACCGGGGTCTCTCGGAAGCAGAGGC 3690
Qy 241 CCTCATCTTCGCCATGCCCCCTGGAGAGCAGCAGCGGTCTCTCATCTCTGCCACCTCAAG 120
Db |||||||
3691 CCTCATCTTCGCCATGCCCCCTGGAGAGCAGCAGCGGTCTCTCATCTCTGCCACCTCAAG 3570
Qy 301 TGAAGCCCGTGGAATCTTACGGCCCTGTGACCTACATTTGTCAGTGCAGCTAGAAGGC 360
Db |||||||
3751 TGAAGCCCGTGGAATCTTACGGCCCTGTGACCTACATTTGTCAGTGCAGCTAGAAGGC 3810
Qy 361 GCAGCTGACACACACTGCGCTCCGACATCTTTGACTGCTGCTACTGACCAAGCAAGTC 420
Db |||||||
3811 GCAGCTGACACACACTGCGCTCCGACATCTTTGACTGCTGCTACTGACCAAGCAAGTC 3870
Qy 421 TCCGGGGTGGCACTTACACCTTCGGCAGCGCATGTGTACGAAGCAGGAATGGGTCCC 480
Db |||||||
3871 TCCGGGGTGGCACTTACACCTTCGGCAGCGCATGTGTACGAAGCAGGAATGGGTCCC 3930
Qy 481 TACAGAGCCCTCGGACAAAGTCTCTTGGAGGGCCAGCACTGGCTCTGAGGAG 540
Db |||||||
3931 TACAGAGCCCTCGGACAAAGTCTCTTGGAGGGCCAGCACTGGCTCTGAGGAG 3990
Qy 541 GAGAGCAGGGGGCGTTCAGCCCAACCTGCGCCAGCACAAGAGCTTCGCAATTCAGACA 600
Db |||||||
3991 GAGAGCAGGGGGCGTTCAGCCCAACCTGCGCCAGCACAAGAGCTTCGCAATTCAGACA 4050
Qy 601 CAGATCCAGAGGGCGGTTCAGCTGTGTGGGCAATCTGGGAGAGGCCAGCGGGCG 660
Db |||||||
4051 CAGATCCAGAGGGCGGTTCAGCTGTGTGGGCAATCTGGGAGAGGCCAGCGGGCG 4110
Qy 661 GCGCTGGCGCCCAAGATCATCCCTACCAACCCAGGACAGACAGCTGTCTGCGGAA 720
Db |||||||
4111 GCGCTGGCGCCCAAGATCATCCCTACCAACCCAGGACAGACAGCTGTCTGCGGAA 4170
Qy 721 TACAGAGCCCTCAAGGGCTCGGCCACCGCACCTTGGCCAGCTGCAAGCAGCTACTC 780
Db |||||||
4171 TACAGAGCCCTCAAGGGCTCGGCCACCGCACCTTGGCCAGCTGCAAGCAGCTACTC 4230
Qy 781 AGCCCCCGGACCTGGTGTCTATCTTGAAGTGTCTTGGGCCCGAGCTGCTCCCTGC 840
Db |||||||
4231 AGCCCCCGGACCTGGTGTCTATCTTGAAGTGTCTTGGGCCCGAGCTGCTCCCTGC 4290
Qy 841 CTGGCCGAGAGGGCTCTCTACTCAGATTCAGAGTGTGAGTACTGCTGGCAGATTTG 900
Db |||||||
4291 CTGGCCGAGAGGGCTCTCTACTCAGATTCAGAGTGTGAGTACTGCTGGCAGATTTG 4350
Qy 901 AGTGCACCCAGTACCTGCACAAACAGCAATCTGCACTGGACCTGAGTCCGAGAAC 960
Db |||||||
4351 AGTGCACCCAGTACCTGCACAAACAGCAATCTGCACTGGACCTGAGTCCGAGAAC 4410
Qy 961 ATGATCATCACCGAATACAACTGCTAAGTCTGGACCT 1001
Db |||||||
4411 ATGATCATCACCGAATACAACTGCTAAGTCTGGACCT 4451

RESULT 7
AR255867
LOCUS AR255867 5207 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6482624.
ACCESSION AR255867
VERSION AR255867.1 GI:27305065
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 5207)
AUTHORS Wei,M.-H., Ketchum,K.A., Di Francesco,V. and Beasley,E.M.
TITLE Isolated human kinase proteins, nucleic acid molecules encoding human kinase proteins; and uses thereof
JOURNAL Patent: US 6482624-A 1 19-NOV-2002;
FEATURES
Location/Qualifiers
source 1..5207
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 99.8%; Score 999.4; DB 6; Length 5207;
Best Local Similarity 99.9%; Pred. No. 8.7e-156;
Matches 1000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GGCAGTCAGTGACACTGGCTTCAGCGGTGTACGCCAGCAGCTGCCAGGCCACTGG 60
Db 3628 GGCAGTCAGTGACACTGGCTTCAGCGGTGTACGCCAGCAGCTGCCAGGCCACTGG 3687
Qy 61 AGCAAGACGAGGCCCTTCGAGAGCAGCAGCGGTGTCTCATCTCTGCACTCAAG 120
Db 3688 AGCAAGACGAGGCCCTTCGAGAGCAGCAGCGGTGTCTCATCTCTGCACTCAAG 3747
Qy 121 AACTTCAGCTTCTGACCAATCTCTGGTGTGGTGTAGGACCTGGTGTGTACCTGC 180
Db 3748 AACTTCAGCTTCTGACCAATCTCTGGTGTGGTGTAGGACCTGGTGTGTACCTGC 3807
Qy 181 AGCGTCAGCAATGCGCTGGGACAGTGACCAACCGGGCTCTCCGGAAGCAGAGCGC 240
Db 3808 AGCGTCAGCAATGCGCTGGGACAGTGACCAACCGGGCTCTCCGGAAGCAGAGCGC 3867
Qy 241 CCTCATCTTCGCCATTCGCCGATTCGGGAGGTGTACGCGATGGGTGTCTGGTC 300
Db 3868 CCTCATCTTCGCCATTCGCCGATTCGGGAGGTGTACGCGATGGGTGTCTGGTC 3927
Qy 301 TGAAGCCCTGGAATCTTACGCGCTGTACCTACATTTGTCAGTGCAGCTTAGAGGC 360
Db 3928 TGAAGCCCTGGAATCTTACGCGCTGTACCTACATTTGTCAGTGCAGCTTAGAGGC 3987
Qy 361 GGCAGCTGGAACCACTGGCTCCGACATCTTTGACTGTCTGTCTGACCCAGCAAGCTC 420
Db 3988 GGCAGCTGGAACCACTGGCTCCGACATCTTTGACTGTCTGTCTGACCCAGCAAGCTC 4047
Qy 421 TCCGGGGTGGCACTTACACCTTCCGACCGCATGTGTAGCAAGCAGCAAGATGGGTCC 480
Db 4048 TCCGGGGTGGCACTTACACCTTCCGACCGCATGTGTAGCAAGCAGCAAGATGGGTCC 4107
Qy 481 TACAGCAGCCCTTCGGAGCAAGTCTCTGGGAGGGCCAGCCACCTGGCTCTGAGGAG 540
Db 4108 TACAGCAGCCCTTCGGAGCAAGTCTCTGGGAGGGCCAGCCACCTGGCTCTGAGGAG 4167
Qy 541 GAGAGCAGGGGGGTGAGCCCAACCTTCGCCAGCAAGAACCTTTCGATTTCCAGACA 600
Db 4168 GAGAGCAGGGGGGTGAGCCCAACCTTCGCCAGCAAGAACCTTTCGATTTCCAGACA 4227
Qy 601 CAGATCCAGAGGGGGCGCTTACGCTGTGTGGCGCATGTCTGGGAGAGGCCAGCGGGCG 660
Db 4228 CAGATCCAGAGGGGGCGCTTACGCTGTGTGGCGCATGTCTGGGAGAGGCCAGCGGGCG 4287
Qy 661 GCGCTGGCGCCCAAGATCATCTCCCTACCAACCCAGGACAGACAGCTGTCTGCGCGAA 720
Db |||||||
4288 GCGCTGGCGCCCAAGATCATCTCCCTACCAACCCAGGACAGACAGCTGTCTGCGCGAA 4347

Qy	721	TACGAGGCCCTCAAGGGCGCTGGCCACCCGGACCTGGCCCCAGCTGCACGAGCCTACCTC	780
Db	4348	TACGAGGCCCTCAAGGGCGCTGGCCACCCGGACCTGGCCCCAGCTGCACGAGCCTACCTC	4407
Qy	781	AGCCCCCGGCACCTGGTGTCTCATCTTGGAGCTGTGTCTCTGGGCCCGAGCTGTCTCCCTGTC	840
Db	4408	AGCCCCCGGCACCTGGTGTCTCATCTTGGAGCTGTGTCTCTGGGCCCGAGCTGTCTCCCTGTC	4467
Qy	841	CTGGCCGAGAGGGGCTCCTTACTCAGAAATCCGAGGTGAAGGACTTACTGTGTGGCAGATTTTG	900
Db	4468	CTGGCCGAGAGGGGCTCCTTACTCAGAAATCTGAGGTGAAGGACTTACTGTGTGGCAGATTTTG	4527
Qy	901	AGTGCCACCCAGTACTGTGCACACCCAGCACATCTCTGCACCTGGACCTGAGGTCCGAGAAC	960
Db	4528	AGTGCCACCCAGTACTGTGCACACCCAGCACATCTCTGCACCTGAGGTCCGAGAAC	4587
Qy	961	ATGATCATCACCGAATAACAACCTGCTCAAGGTGTGGACCT	1001
Db	4588	ATGATCATCACCGAATAACAACCTGCTCAAGGTGTGGACCT	4628

```

RESULT 8
AR442826
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN

AR442826
Sequence 1 from patent US 6670164.
AR442826
AR442826.1 GI:42670270
.
Unknown.
SOURCE
Unknown.
ORGANISM
Unclassified.
REFERENCE
1 (bases 1 to 5207)
AUTHORS
Wei, M.-H., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.
TITLE
Isolated human kinase proteins
JOURNAL
Patent: US 6670164-A 1 30-DEC-2003;
Location/Qualifiers
1. 5207
/organisms="unknown"
/mol_type="genomic DNA"

```

	Query Match	99.8%	Score 999.4;	DB 6;	Length 5207;
	Best Local Similarity 99.9%;	Pred. No. 8,7e-156;			
	Matches 1000; Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;	
Qy	1	GGCCAGTCAGTGACACTGGCCTGCCAGGTGTCAGCCAGCGACGTGCCCAAGGCCACCCTGG	60		
Dd	3628	GGCCAGTCAGTGACACTGGCCTGCCAGGTGTAGCCAGCCAGCTGCCCCAGGCCACCCTGG	3687		
Qy	61	AGCAAAGACGGAGGCCCCCTTGGAGAGCAGCAGCGCGTGTCTCATCTCTGCACACCCCTCAAG	120		
Dd	3688	AGCAAAGACGGAGGCCCCCTTGGAGAGCAGCAGCGCGTGTCTCATCTCTGCCACCCCTCAAG	3747		
Qy	121	AAC TTCAGCTTTGACCATCTCTGGTGGTGGTGAGGACCTGGGTGTGTATACCTGTC	180		
Dd	3748	AAC TTCAGCTTTGACCATCTCTGGTGGTGGTGAGGACCTGGGTGTGTATACCTGTC	3807		
Qy	181	AGCGTGAGCAATGCGCTGGGGGACAGTGACACACACCGGGCGTCTCTCCGGAAGGCAGAGCGC	240		
Dd	3808	AGCGTGAGCAATGCGCTGGGGACAGTGACACACACCGGGCGTCTCTCCGGAAGGCAGAGCGC	3867		
Qy	241	CCCTCATCTTCGGCATGCCCGBATATCGGGGAGGTGTACCGGANATGGGGTGTCTGTGTC	300		
Dd	3868	CCCTCATCTTCGGCATGCCCGBATATCGGGGAGGTGTACCGGANATGGGGTGTCTGTGTC	3927		
Qy	301	TGGAAGGCCCGTGGAAATCCTACGGGCCCTGTGACCTACATTGTGCAGTGCACGCTTAGAAGGC	360		
Dd	3928	TGGAAGGCCCGTGGAAATCCTACGGGCCCTGTGACCTACATTGTGCAGTGCACGCTTAGAAGGC	3987		
Qy	361	GGCAGCTGGACCAACACATGGGCTCCGACATCTTTGATCTGTCTACTCTGACCAAGCAAGCTC	420		
Dd	3988	GGCAGCTGGACCAACATGGGCTCCGACATCTTTGATCTGTCTACTCTGACCAAGCAAGCTC	4047		

Qy	421	TC	CGGGGTGGCACCTTACACTTCCGCACGGCATGTGTCAAGAGCAGGAATGGGTCCC	480
Db	4048	TC	CGGGGTGGCACCTTACACTTCCGCACGGCATGTGTCAAGAGCAGGAATGGGTCCC	4107
Qy	481	TAC	AGCAGCCCCCTCGGAGCAAGTCCTCTCGGAGGGCCACGACCTGGGCTCTGAGGAG	540
Db	4108	TAC	AGCAGCCCCCTCGGAGCAAGTCCTCTCGGAGGGCCACGACCTGGGCTCTGAGGAG	4167
Qy	541	GAG	AGCCAGGGGGGGTTCAGGCCAACCCCTGCCAGACAAAGACCTTTCGATTTCCAGACA	600
Db	4168	GAG	AGCCAGGGGGGGTTCAGGCCAACCCCTGCCAGACAAAGACCTTTCGATTTCCAGACA	4227
Qy	601	CAG	ATCAGAGGGGCGCTTCAGGGTGTGGGGCAATGCTGGAGAGGCCACGGGGCGG	660
Db	4228	CAG	ATCAGAGGGGCGCTTCAGGGTGTGGGGCAATGCTGGAGAGGCCACGGGGCGG	4287
Qy	661	GGC	TGGCGCCCAAGATCATCCCTACCAACCCCAAGGCAAGACAGCAGTGCTGCCGGA	720
Db	4288	GGC	TGGCGCCCAAGATCATCCCTACCAACCCCAAGGCAAGACAGCAGTGCTGCCGGA	4347
Qy	721	TAC	GAGGCCCTCAAGGGCTGCGCCACCCGCACTGGGCCAGCTGCAACGAGCCTACCTC	780
Db	4348	TAC	GAGGCCCTCAAGGGCTGCGCCACCCGCACTGGGCCAGCTGCAACGAGCCTACCTC	4407
Qy	781	AG	CCCCCGGCACTGTGTGTCTATTGGAGCTGTGTCTTGGGCCGAGCTGCTCCCTGC	840
Db	4408	AG	CCCCCGGCACTGTGTGTCTATTGGAGCTGTGTCTTGGGCCGAGCTGCTCCCTGC	4467
Qy	841	CT	GGCCGAGGGGCTCCTACTCAGATCCGAGGTGAGACTTACCTGTGGCAGATGTTG	900
Db	4468	CT	GGCCGAGGGGCTCCTACTCAGATCCGAGGTGAGACTTACCTGTGGCAGATGTTG	4527
Qy	901	AGT	GCCACCCAGTACCTGCAACACGACACATCTGCACCTGGACCTTGAGTCCGAGAAC	960
Db	4528	AGT	GCCACCCAGTACCTGCAACACGACACATCTGCACCTGGACCTTGAGTCCGAGAAC	4587
Qy	961	AT	GATCATCACCGAATAACACCTGCTCAAGGTCGTGGAACCT	1001
Db	4588	AT	GATCATCACCGAATAACACCTGCTCAAGGTCGTGGAACCT	4628

```

RESULT 9
AR442827
LOCUS
DEFINITION
ACCESSION
AR442827
VERSION
AR442827.1 GI:42670271
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 5207)
AUTHORS
Wei, M.-H., Ketchum, K.A., Di Francesco, V. and Beasley, E.M.
TITLE
Isolated human kinase proteins
JOURNAL
Patent: US 6670164-A 3 30-DEC-2003;
LOCATION/Qualifiers
1. .5207
source
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN

```

Query Match	99.8%;	Score 999.4;	DB 6;	Length 5207;
Best Local Similarity	99.9%;	Pred. No. 8.7e-156;		
Matches 1000;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	GGCCAGTCACTGACCTGGCTGCCAGGTGTTCAGCCAGCCAGCTGCCAGGCCACCTGG	60	
Db	3628	GGCCAGTCACTGACCTGGCTGCCAGGTGTTCAGCCAGCCAGCTGCCAGGCCACCTGG	3687	
Qy	61	AGCAAGACGGAGCCCCCTGGAGAGCAGGACCGGTGCTCTATCTCTGCCACCCCTCAAG	120	
Db	3688	AGCAAGACGGAGCCCCCTGGAGAGCAGGACCGGTGCTCTATCTCTGCCACCCCTCAAG	3747	

QY 185 TGAGCAATGCGCTGGGACAGTGACCAACCGGGCGTCTCTCGGAAGCAGAGCGCCCT 244
 Db 17287 TGAGCAATGCGCTGGGACAGTGACCAACCGGGCGTCTCTCGGAAGCAGAGCGCCCT 17346
 QY 245 CATCTTCCCATGCGCCGATATCGGGAGGTGTACCGGATGGGTGCTGCTGGTCTGGA 304
 Db 17347 CATCTTCCCATGCGCCGATATCGGGAGGTGTACCGGATGGGTGCTGCTGGTCTGGA 17406
 QY 305 AGCCCGTGGAAATCTTACCGCCCTGTGACCTATCTGTCAGTGACGACCTAGAGCGGCA 364
 Db 17407 AGCCCGTGGAAATCTTACCGCCCTGTGACCTATCTGTCAGTGACGACCTAGAGCGGCA 17466
 QY 365 GCTGGACCACTGGCCCTCCGACATCTTTGACTGCTGCTACTGACCAAGAGCTCTCCC 424
 Db 17467 GCTGGACCACTGGCCCTCCGACATCTTTGACTGCTGCTACTGACCAAGAGCTCTCCC 17526
 QY 425 GGGTGGGACCTACACCTTCCGACGCGATGTGTGACGACGAGGAATGGTCCCTACA 484
 Db 17527 GGGTGGGACCTACACCTTCCGACGCGATGTGTGACGACGAGGAATGGTCCCTACA 17586
 QY 485 CGAGCCCTCGGAGCAAGTCTCTTGGAGGGCCAGCCACTGGCCCTCTGAGGAGGAGA 544
 Db 17587 CGAGCCCTCGGAGCAAGTCTCTTGGAGGGCCAGCCACTGGCCCTCTGAGGAGGAGA 17616
 QY 545 GCCAGGGCGGTGAGCCCAACCCCTGCCAGCAGCAAGACCTTCGCATTCAGACACAGA 604
 Db 17617 GGAATGGAGGGGACCAACCCAGCCCTGACCGCTGTCCCTTGA----- 17661
 QY 605 TCCAGAGGGCGGCTTACGCTGGTGGCGCAATGCTGGGAGAGCCAGCGGGCGGCGC 664
 Db 17662 ---GCAGGGCGGCTTACGCTGGTGGCGCAATGCTGGGAGAGCCAGCGGGCGGCGC 17718
 QY 665 TGGCCGCGCAAGATCATCCCTACCAACCCAGGACAGACAGTGTGCGCGAATACG 724
 Db 17719 TGGCCGCGCAAGATCATCCCTACCAACCCAGGACAGACAGTGTGCGCGAATACG 17778
 QY 725 AGGCCCTCAAGGGCTGGCCACCGGACCTGGCCAGCTGCAAGCAGCTACCTCAGCC 784
 Db 17779 AGGCCCTCAAGGGCTGGCCACCGGACCTGGCCAGCTGCAAGCAGCTACCTCAGCC 17838
 QY 785 CCGGACCTGGTGTCTATCTTGGAGTGTGCTCTGGGCGGAGCTGCTCCCTGCGCTGG 844
 Db 17839 CCGGACCTGGTGTCTATCTTGGAGTGTGCTCTGGGCGGAGCTGCTCCCTGCGCTGG 17898
 QY 845 CCGAGAGGGCTCTCTACTCAGAAATCCGAGGTGAAGACTTACCTGTGGCAGATGTTAGTG 904
 Db 17899 CCGAGAGGGCTCTCTACTCAGAAATCCGAGGTGAAGACTTACCTGTGGCAGATGTTAGTG 17958
 QY 905 CCACCCAGTACTGCAACACGACGACATCTTGCACCTGGACCTGAGTCCGAGAACATGA 964
 Db 17959 CCACCCAGTACTGCAACACGACGACATCTTGCACCTGGACCTGAGTCCGAGAACATGA 18018
 QY 965 TCATCACCAGAAATACACCTGCTCAAGGTGCTGGACCT 1001
 Db 18019 TCATCACCAGAAATACACCTGCTCAAGGTGCTGGACCT 18055
 RESULT 14
 AY603754
 LOCUS Mus musculus Mus musculus 4740 bp mRNA linear ROD 01-AUG-2004
 DEFINITION Mus musculus obscurin-MLCK mRNA, complete cds.
 ACCESSION AY603754
 VERSION AY603754.1 GI:50660397
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 4740)
 Sutter, S.B., Raeker, M.O., Borisov, A.B. and Russell, M.W.
 Orthologous relationship of obscurin and Unc-89: phylogeny of a
 novel family of tandem myosin light chain kinases

JOURNAL Dev. Genes Evol. 214 (7), 352-359 (2004)
 PUBMED 15185077
 REFERENCE 2 (bases 1 to 4740)
 AUTHORS Sutter, S.B., Raeker, M.O. and Russell, M.W.
 TITLE Direct Submission
 JOURNAL Submitted (21-APR-2004) Pediatric Cardiology, University of
 Michigan, 1500 E. Medical Center Dr., Ann Arbor, MI 48109, USA
 FEATURES Location/Qualifiers
 source
 1..4740
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 CDS
 1..4740
 /codon_start=1
 /product="obscurin-MLCK"
 /protein_id="AA080900.1"
 /db_xref="GI:50660398"
 /translation="MOVITIEDVQVQDMAQFDVIEGHPPPIVTWYKSGTQLTSSAR
 LSOQDCTTVSLVLTVDVPHDAGVYTCVANNAGGOVLCKAEHLVHGDKLDAENQVYR
 RKLHSFDVBEIEGRGVFGVFRVQVHKGKMFCAAKFIPLSKTRAQYQBRDLATL
 GHLVTLGLOQFETRKTLILILELCSSEELDLRFKGVVTEAEVKYIQQLVGSLHY
 LHSGLIHLDDIKPNNILMHPAREDIKICDGFQAKITPSEPQSKYSGVFPVPEII
 EQNPVSGSDIWMGVISYLSLTCSFPAGESDRATLLNVLEGVSSPFAHLSSED
 AKDQIKATLORTPOARSTQCLAHWPFLKMPAEAEHFIKTKLKLARWQSL
 MSYKILVMSIPELLOGPDSPLGVARHLRGEASGASSSSSSDNELAPFAKSL
 PPSVTHSPLHPRGFLRPSASLPEETEASHTADAAVPASQSPQAGPPASGCVPRHS
 VISLFFYQAGGAEKNGTSGAKRHPARRHLHLKGGYIARALPGLRPLMEYLSLEE
 EAAEEQASLMTKTPFETALRLPSSVREVPVPSGSHLNDNPVTTGSPGACKBQLLF
 PPSTGLTHETAKRQHEGFLQBSVFPFPGSGDRSGKQEGSQDSGRGPAKSGHS
 ELGSGGEGGPPSSQSLGSLPQSLKELSTSCPLPSEQPAAPPTVPSPLGSE
 KEDQDGLSGGPPVPSGASQVDSLDTGELSEAGDTCDTTPPQROEQATTR
 KFSLESRGVAGVAGYCTPAGGAGGMLGQGLPWARMANAVSSGSEEDANATESPQ
 PLESGLPIASGVPLTSPSLTPWEVEVQSLVQIRDLSDGDAADTISLDSIVDP
 AYNLSDLYDIKYLFPFEMI PRVPKPIEQPESGSETEAGQGLADFLLEAAWMPGE
 LGRAGLEITEPEEPDGLLEAAVGRKRKWSRSLGFPQFGRCLSGSEPEVLGL
 RORVKAEMHSIRILKGRPEGPERGPRKAGLSRLSLGKLRDQSLDSAEVVLGO
 SVTLACVLAQPTAQTMSKDVLLSESHLLISLTKNFOLLTLVVKVEDLGTTC
 CYNPLGTAVTTGVLRKAERSSPREVEGELYKADVLLVWKPVESGCVTYIYQCCI
 EBSGWTTLASDISDCYLTGKLSRGMYIFRTACVSKAGMGPPSPSEQVLLGGPNHL
 ASBESRGRPAQLLPSTKTFQFQIRGRFVVRQCRKASGRALAAKATVPYQPED
 KVALREYEAALKRLHPLHQAALYLSRHLVILRCLSGPELLPSLAERESYSED
 VKDYLMWLSATQYVLAHQHILHLDLRSENMMVBYNLLKVIDLGNQSLDQEKVPAPE
 NKDYLETMAPELLEGGAGVPTDIIWAIGVTAFTMLSGVPESEGTGRLDKQGRKGL
 IRLSRVAGLSGGNAFLQSLCAQPWGRPCASTCLOGGMLTEGPTGSRTPVTFPT
 VRLRAFVREKRRALLYKKNLAQVR"
 439..1206
 /note="Region: serine/threonine protein kinase domain"
 4015..4725
 /note="Region: serine/threonine protein kinase domain"
 ORIGIN
 Query Match 73.8%; Score 738.4; DB 10; Length 4740;
 Best Local Similarity 84.2%; Pred. No. 1.3e-112;
 Matches 845; Conservative 0; Mismatches 156; Indels 3; Gaps 1;
 QY 1 GGCAGTCAGTGAACACTGGCCCTGCCAGGTGTACCCAGCCAGCCAGCTGG 60
 Db 3259 GGCCAATCAGTGAACACTGGCCCTGCCAGGTGTGTGGCCAGCACTGGCC 3318
 QY 61 AGCAAGACGAGCCGCCCTGGAGAGCAGCAGCGTGTCTCATCTCTGCCACCTCAAG 120
 Db 3319 AGCAAGATGGGTCCTCTGGAGAGCAGCGCCAGCTCTCTCATCTCTCCACCTGAAG 3378
 QY 121 AACTTCCAGCTTCTGACCATCTCTGGTGGTGTGTGAGGACCTGGGTGTGTACACCTGC 180
 Db 3379 AACTTCCAGCTTCTGACCATCTCTGGTGAAGAGGAGGATCTGGGCACATATACCTGC 3438
 QY 181 AGCTGAGCAATGCGCTGGGACAGTGAACCAACCGGCGCTCTCCGGAAGAGCAGCGC 240
 Db 3439 TGTGTGACCAACCGCTGGGACAGCAGTCAACACAGGTGTCTCCCGGAAAGCAGAGCGC 3498
 QY 241 CCCTCATCTTCGCGATCCCGGATATCGGGAGGTGTACCGCGATGGGTCTCTGCTGTC 300


```
Db 3499 CTTTCATCTTTCCACGCCCTGAGGTGGGGAACTATACAAAGGATGGGTGCTGCTGGTC 3558
Qy 301 TGAAGAGCCCGTGAATCTACGACCCCTGTGACCTACATTTGTGAGTGCAGTGCAGCTTAGAGGC 360
Db 3559 TGGAGCCTGTGAAATCTGTGCGCCGCTGACCTACATTTGTGAGTGCAGTGCAGCTTAGAGGAG 3618
Qy 361 GGCAGCTGGACACACTGGCTCCGACATCTTTGACTGTGCTGCTACCTGACCAAGCAAGCTC 420
Db 3619 GGCAGCTGGACAAACCCCTGGCTTCAGACATCTCTGACTGTGCTACCTGCTACCTGCAAGCTA 3678
Qy 421 TCCCGGGGTGGACCTACACCTTCGACCGCATGTGTAGAGGAGGAGCAAGATGGGTCCC 480
Db 3679 TCGAGGGGTGGCATGTATATCTTCGCGACAGCATGTGTGAGCAAGCAGGAATGGGCCCC 3738
Qy 481 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGGCCGACGACCTGCGCTCTGAGGAG 540
Db 3739 TACAGCAGCCCTCAGAACAGTCTCTCTCGGAGGACCCCAACCTGCGCTCTGAGGAG 3798
Qy 541 GAG---AGCCAGGGCGGTGAGCCCAACCCCTGCGCAGCACAAAGACCTTGGCATTCAG 597
Db 3799 GAAAGCAGCGGGGGAGGCAGGCCAGCTTCTCTCCAGCACAAAGACCTTTGCCCTTCCAG 3858
Qy 598 ACACAGATCCAGAGGGCGCTTACGCGTGTGCGGATGCTGGAGAGGCGCAGCGGG 657
Db 3859 ATGAGATCCGAGGGGCGCTTACGTTGTGCGGAGTGCAGGAGAGGCAAGTGGG 3918
Qy 658 CGGGCGCTGGCCGACGATCATCCCTACACCCGACGACCAAGGACAGACAGCAGTGTGCGC 717
Db 3919 CGGGCGCTAGCTAGATCGTTACCATCTGCTTACCAACCTGAGGACAGACAGCTGTACTAGA 3978
Qy 718 GAATACGAGGCTTCAAGGGCTGCGCCACCGCAGCTGCGCCAGCTGACGACGCTTAC 777
Db 3979 GAATACGAGGCTTAAAGAGACTGCACACCCACCATCTGGCCCAACTCCATGCGCGCTAC 4038
Qy 778 CTCAGCCCCCGGACCTGCTCATCTTGGAGCTGTCTCTGAGCTGTGCTCTGCGCCCTGAGCTGCTACCC 837
Db 4039 CTCAGTCCCGGACCTGCTCATCTTGGAGCTGTCTCTGAGCTGTGCTCTGCGCCCTGAGCTGCTACCC 4098
Qy 838 TCCCTGGCGGAGAGGCGCTCTTACTCAGAACTCGAGGTGAAGGATCACTGTGGCAGATG 897
Db 4099 TCTTTGGCAGAGAGGGAATCTTACTCAGAGTCTGATGTGAAGGACTACCTGTGGCAGATG 4158
Qy 898 TTGAGTGCCACCCAGTACTGCACACAGCAGACATCTTGCACCTGAGCTGAGCTCGAG 957
Db 4159 CTGAGTGCCACCCAGTACTGATGCCCAACACATCTCTGCGACCTGAGCTGAGCTCGAG 4218
Qy 958 AACATGATCATCACCAGAAATACAACTGCTCAAGGTCGTGGACCT 1001
Db 4219 AACATGATGCTCACTGAGTACAACTGCTTAAGGTATAGACCT 4262

RESULT 15
BC046431 1183 bp mRNA linear ROD 16-SEP-2003
LOCUS Mus musculus cDNA sequence BC046431, mRNA (cDNA clone
IMAGE:4006702), partial cds.
ACCESSION BC046431
VERSION BC046431.1 GI:28386252
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1183)
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, J.D., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
```

```
TITLE NIH-MGC Project URL: http://mgc.nci.nih.gov
JOURNAL Contact: MGC help desk
MEDLINE Email: cgaabs-rc@mail.nih.gov
PUBMED Tissue Procurement: Gilbert Smith, Ph.D.
REFERENCE cDNA Library Preparation: Life Technologies, Inc.
AUTHORS cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
TITLE DNA Sequencing by: Baylor College of Medicine Human Genome
JOURNAL Sequencing Center
REMARK Center code: BCM-HGSC
COMMENT Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Lounseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 92 Row: d Column: 18.
FEATURES
source
1..1183
/organism="Mus musculus"
/mol_type="mRNA"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:4006702"
/tissue_type="Mammary tumor metastasized to lung.
MMTV-LTR/Wnt1 model. Expression driven by an MMTV-LTR
enhancer."
/clone_lib="NCI CGAP_Lu30"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..1183
/clone="BC046431"
/db_xref="LocusID:216790"
/db_xref="MGI:2664530"
<1..1027
/clone="BC046431"
/codon_start=2
/product="BC046431 protein"
/protein_id="AAH46431.1"
/db_xref="GI:28386253"
/db_xref="LocusID:216790"
/db_xref="MGI:2664530"
/translation="VSKAGMGYPSPSEOVLLGGPNHLASEESRGPALLPTKTK
FAFMQIPIRGFSVVRQCRKASRALAAKIVPQPEKTVLRREYKALKRLHHPHLA
QLHAYLSPHRLLELDCSGPELLPSLAERESYSDVDKYLQMLSATQYLAHOHI
LHLDURSNMVMTEYNLLKVDLGNQAQSLDQKVPAPENFKLIRSCVAGLSGGAFLQS
PQTDTWAIQVTAFLMSGEYPSSEGTDLQKLRKGLIRLRCVAGLSGGAFLQS
SLCQWGPGRPCASTCLOCGWLTEBPTGSRPTPTVPTVRLRAFVREKRRLYYK
HNLQAVR"
134..892
/misc_feature
/gene="BC046431"
```


THIS PAGE BLANK (USPTO)

/product="obscurin"
/protein_id="CAC44768.1"
/db_xref="GI:15026974"
/db_xref="GOA:Q96AA2"
/db_xref="UniProt/Tc:EMBL:Q96AA2"
/translation="MDQDPFSGAPRLTPKPAVSVYSGKDATLSQIVGNPTPOVSHVE
KQDPQVAGAPRRLAQDGLIPLTILDLHSDGSQVPCARNVAGEAPAAVGLGVQDRAE
AACAQAPHFLLRPTSLIRVREGSEATFRCVGSPPVPCARNVSGKDRGEPDGPGRVVR
SELGEASALRIARPRDGPSTVEVRAENPLGSAASAAALVDSADTAPSTPGSTAA
LLAHLQRREARMAAGPASPSTGTCTTPEXGHARSLCVSYDGEKPTETVMKKDGO
LVTRGHRVYEDQAFENLKIPLFKQSDRGLTCTASNLLAGTQVSSVLVVVREPAVR
FKKLLQDLVREKESATFLCEVPCSTDAAMFKETRLWASAKTGYIEEGETRELTPVR
VNSADDAVYLCETPEGSPVAELAVQGNLKLKLPKTAVRVGTAMPCEVLAVPVGP
VHMLRVBEVVAGGRVAVSABSTGTHLTITSCQLESDVGVQAMGDQGTSTPRFCVAP
RKPFLQPPVDPVVKARMSVLSLWSPPHPGSHVPTIDGILVEKKGLGTVYTRICHSEA
EWVAPELTVDVAENAGFWRVLDQALKASGVYEIEFTDTVHLKPLAVRTPLKVAQVA
VEGEVETVSDVLTVASAGFWRVLDQALKASGVYEIEFTDTVHLKPLAVRTPLKVAQVA
KFAVNGTSSIRMEVRAEPGLTANKRPAAARAEVLAHLHEBAQALAEALSQAAATWL
KDGRTLSPGPKYVQASAGRIKLVREDVARDAGILVCVSRGRILVQLSVQGLAPRL
HKDWAGSCVDVAGGAPQPECTSEAHVHVHYKQMLGHSGERPLQEDVGTGTHRLV
AAATVTRQDEGTVACRYGVEISVDFRLKPSKVPFAKEQLAHKQALQAEASQILSEFV
QAQTEVTYKDGKLLSSSKCMEATGCTRLVVOQAQDAGEYSCEAGQRLSCV
LDVXPKVFAKQDVQAHSEVOAEAGANATLSCVSAQAQAEVVMYKDGKLLSSKLVHY
EAKGRRLVVOQAQTDAGDYSCEAGQRLSCVFLHITPEPMQAEKQSVHNEVQAEA
GASMLSCVEVAQAEVTYVYKDGKLLSSSKCMEATGCTRLVVOQAQDAGEYSCE
BAGQRLSCVFLHITPEPMQAEKQSVHNEVQAEATMLSCVEQAQTEVTYVYKDGK
KLSSSKVRMEVKGCTRLVVOQAQDAGEYSCEAGQRLSCVFLHITPEKAVPAKEQ
LVHNEVTEAGASATLSCVEQAQTEVTYVYKDGKLLSSSKVRMEVKGCTRLVVOQA
QADAEVTYKDGKLLSSKVRMEVKGCTRLVVOQAQDAGEYSCEAGQRLSCVFLHIT
EPKVYFAKEQVHNEVQAEATMLSCVEQAQAEVTYVYKDGKLLSSSKVRMEVKG
CTRLVVOQAQDAGEYSCEAGQRLSCVFLHITPEKAVPAKEAPSAQAEVQAEATSA
TLSCVEQAQAEVTYVYKDGKLLSSSKVRMEVKGCTRLVVOQAQDAGEYSCEAG
QRLSCVFLHITPEKAVPAKEAPSAQAEVQAEATMLSCVEQAQAEVTYVYKDGKLLS
SSKVRMEVKGCTRLVVOQAQDAGEYSCEAGQRLSCVFLHITPEKAVPAKEAPSAQ
PLVKEHETILTLTAPSAATMLVYKDGVEYIRSKRHTLHVAELBQITLSERPCHRE
DSALYPRVGAEGQDPVQVBEVNAFCRLLEPAGVEYCEGELGTVTLACLSFACPAVVR
CGNTQPRVGRFQWAPGPNVSTVLGLRAEDAGEYVCESEDRDHTLSAQLTVSVRVPV
FMGSLTVYVAEGEATFOCVVPSVDVAWFRDGLALPSEKFAISQGSASHLTSIL
DLVLADAGQITVEAGASASAAALVREAAPLVFLKLPQTEVERESSITVELTRPMP
ELRWATNALAPGKNVEIHASGARHLVLHNVGADRGFPFGCTPDDTKQAKLVEM
RQVRLVGLQAVBARQGTATMEVQLSHADVGSMTDGLRFQGGTCHLVAVRGMHT
TILSLGRPDUSGLMFAKGVHTSARLVTLPVFSRPLQDVTTEKVEQITLECLS
RPNVVRLMKDGLVEPAGTMAIAQAACRSITLYRCEFAQDQVYVCDHAQKASFLVQ
KVQGRVITYRVLRAEDAGEITQVFAENASQATLYRCEFAQDQVYVCDHAQKASFLVQ
GVLEQCVSRAAQVWRMGKQSLQPGPKVLELDGLKRLITSDVHAEDSDTLQAD
DKVTSQAQVFEVQESITLYRGLQDQVPAWPAPECTSLPSVPRKMLLGTVLTQAG
GNVLGREGTTHLMALRKTCTMTGPHVTVGKSSSARLVSDIPVLTPLRLEPKTG
RELQSVLSCDFRPAPQVWTKDTPDTPSPSKFMSLEQMAELRLRLPADAGVY
RCQASGASHSETVYVAREVTVPGLQDAEATBEGWAFSCELSHEDDEVEVSLNGMP
LYNDSFHETLHKGRRHTLVQITIQADAGIVRAASLKVSTYSARLVKRVKPVFLKALD
DLASAEQITLQCEVDEPAHVNRKDGQVSDKDPFLHTAGTGLRVLHVHDPSD
AGLYTCHVSGESTRAARVRHQDKLVYTKRLTVMLEVGESCEPVCVLSHESADPAMV
TVGKTVGSSSRFOATVRQDKLVLVREAAPSDAGEVYVFSRGLVTKASLIVREPA
IILPLEQDVAPEEDVELCELSRAGTPVHMLKORKAIRKQKDVDCBGTWAMLVIR
GASLKAGYTVCEVASKHSLVVEKANCPTBELNLQVBERGTATVYCTKEHPAA
VTVWRKGLLELAGSKHQPSEGLTILYLSALEKADSDTYCTGOAQRASQLVQ
RVHYITLEDLVQDGSSEATFRCISPAANTYVHVEPLDTLHANELINDAQPGY
HVLTLRALQKDSGTIFYPAGDQASAAALVTEKPSFRELDTATIGEDTLVCE
TSTCDTPMCTWQKTLRGARCOLSHHGAQLITGATILQDSRGKEAGACSSS
ITVARHVPFOBALKDLVEEGGATLRCVLSVAPKVCYGNVLPRGDKYSLRQ
EGAMLELVNRLPQDSGRYSCSFGDTTSATLVTALPAQITGLUKLEATGEGYEL
LRCELSTKPAVWRKGSSETLRDGRYCURQDQACELOIRGLAWDAAYSVCGVEEL
TSASLRPMPAHTGRUHQESBEGATLQDCELSKAAVPEWRKGSRLSDGRHET
RDQGAVELQICGLAVADAGESYCGEERTSATLVKALPAGTTLRGRNEARVBAET
AMWLCELSKPAVPEWRKGSLENLRDGRYILRQEGTRCELOICGLAMADAGEYLCVGO
ERTSATLTPALPAQVPEVDQNEAREGATLVCELSNSAPVEMRKGSSETLRDGRY
SLRQDGTKELOIRGLAWDAGEYSCVQBERTSAMLTPALPIKTEGRLNEARET
ATAVLRCELSKPAVPEWRKGSLENLRDGRYILRQDGRARCELOIRGLVADAGEYLCMC
GKERTSAMLTPALPAQVPEVDQNEAREGATLQDCELSKAAVPEWRKGSSETLRDGR
RHSRLSDGRCELOIRGLAWDAGEYSCVQBERTSAMLTPALPAQVPEWRKGSSETLR
EGNATVLOCELSKAAVPEWRKGSLENLRDGRYILRQDGRCELOIRGLVADAGEYLC
VQBERTSAMLTPALPAQVPEVDQNEAREGATLQDCELSPTATVWSKGLQLO

[illegible]

ORIGIN

	Query Match	100.0%;	Score 1001;	DB 9;	Length 20435;
	Best Local Similarity	100.0%;	Pred. No. 4.1e-192;		
	Matches 1001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AAAGATGGGAAGAGCTGAGCTCCAGCTTGAAAGTGCATGTAGAGGCCAAAGGCTGCAGA	60		
Db	3120	AAAGATGGGAAGAGCTGAGCTCCAGCTTGAAAGTGCATGTAGAGGCCAAAGGCTGCAGA	3179		
Qy	61	CGAGAGCTGGTGGTGAGACGAGCGAGCGACAAGACGGATGCCGGGGACTACAGCTGCGAGGCC	120		
Db	3180	CGAGAGCTGGTGGTGAGCAGGCGAGCGACAAGACGGATGCCGGGGACTACAGCTGCGAGGCC	3239		
Qy	121	AGGGGGCCAGAGGGGTCTCTTTCCGCTTGCACTCACAGAGGCCAAAGATGATGTTTGCAAAG	180		
Db	3240	AGGGGGCCAGAGGGGTCTCTTTCCGCTTGCACTCACAGAGGCCAAAGATGATGTTTGCAAAG	3299		
Qy	181	GACAGTCAATGATCAATGAGGTGACGCTGAGCGGGGGCCAGTGCCTAGCTGAGCTGT	240		
Db	3300	GACAGTCAATGATCAATGAGGTGACGCTGAGCGGGGGCCAGTGCCTAGCTGAGCTGT	3359		
Qy	241	GAGGTGCCCCAGGCCCCAGA	CGAGGTGACGTGTACAGGATGGGAAGAGCTGAGCTCC	300	
Db	3360	GAGGTGCCCCAGGCCCCAGA	CGAGGTGACGTGTACAGGATGGGAAGAGCTGAGCTCC	3419	

QY	301	AGCTCAAAAGTGGGATGGAGTCAAGGGTGCACACGAGGCTGGTGTCCACAGGGG	360
Db	3420	AGCTCAAAAGTGGGATGGAGTCAAGGGTGCACACGAGGCTGGTGTCCACAGGGG	3479
QY	361	GCCTAAGCAGATGCTGGGAGTACAGCTGTGAGGCTGGGGGCACAGAGTCTCTTCCAC	420
Db	3480	GCCTAAGCAGATGCTGGGAGTACAGCTGTGAGGCTGGGGGCACAGAGTCTCTTCCAC	3539
QY	421	CTGCATCAACAGCCCAAGGGGTGTTTGCAGAGGAGCAGTCAAGTGCATTAATGAGGTG	480
Db	3540	CTGCATCAACAGCCCAAGGGGTGTTTGCAGAGGAGCAGTCAAGTGCATTAATGAGGTG	3599
QY	481	CAGGCTGAGGGGGACCACTGCTGAGCTGTGAGGTCGGCCACGCCCCAGACGGAG	540
Db	3600	CAGGCTGAGGGGGACCACTGCTGAGCTGTGAGGTCGGCCACGCCCCAGACGGAG	3659
QY	541	GTGAGCTGTGTAACAGGACGGGAAGAGTGTGAGCTCCAGCTCAAAAGTACGATGAGGTC	600
Db	3660	GTGAGCTGTGTAACAGGACGGGAAGAGTGTGAGCTCCAGCTCAAAAGTACGATGAGGTC	3719
QY	601	AAGGCTGCACACGAAGCTGTGTAGTCAGCAGGTGGGCAAGCAGATGCTGGGAGTAC	660
Db	3720	AAGGCTGCACACGAAGCTGTGTAGTCAGCAGGTGGGCAAGCAGATGCTGGGAGTAC	3779
QY	661	AGCTGCGAGGCTGGGGCCAGAGAGTCTCTTTCAACTGCACATCACAGAGCCCAAGGCA	720
Db	3780	AGCTGCGAGGCTGGGGCCAGAGAGTCTCTTTCAACTGCACATCACAGAGCCCAAGGCA	3839
QY	721	GTGTTTGCAGAGGACAGTGTGTGATTAATGAGTGCAGAGTGCAGGAGGGGCCAGTGCC	780
Db	3840	GTGTTTGCAGAGGACAGTGTGTGATTAATGAGTGCAGAGTGCAGGAGGGGCCAGTGCC	3899
QY	781	ACACTGAGCTGTGAGGTGCCCGCCAGACAGAGGTGACGTGTAACAAGGATGGGAAG	840
Db	3900	ACACTGAGCTGTGAGGTGCCCGCCAGACAGAGGTGACGTGTAACAAGGATGGGAAG	3959
QY	841	AAGCTGAGCTCCAGTTCGAAAGTGCAGATAGGCTGCGGGTGCATGCGGAGCTGGTG	900
Db	3960	AAGCTGAGCTCCAGTTCGAAAGTGCAGATAGGCTGCGGGTGCATGCGGAGCTGGTG	4019
QY	901	GTGCAGAGGAGGCGCAGCAGATGCTGGGAGTACACTGTGAGGCTGGGGGCCAGCGG	960
Db	4020	GTGCAGAGGAGGCGCAGCAGATGCTGGGAGTACACTGTGAGGCTGGGGGCCAGCGG	4079
QY	961	CTCTCTTCCACCTGGATGTTTCAGAGCCCAAGGCGGTGT	1001
Db	4080	CTCTCTTCCACCTGGATGTTTCAGAGCCCAAGGCGGTGT	4120
RESULT 2			
LOCUS	CQ730656 18524 bp DNA linear PAT 03-FEB-2004		
DEFINITION	Sequence 16590 from Patent WO02068579.		
ACCESSION	CQ730656		
VERSION	CQ730656.1 GI:42305092		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.		
TITLE	Kits, such as nucleic acid arrays, comprising a majority of human exons or transcripts, for detecting expression and other uses thereof		
JOURNAL	Patent: WO 02068579-A 16590 06-SEP-2002;		
FEATURES	PE Corporation (NY) (US)		
source	1.18524		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			

Query Match		77.5%;	Score 775.8;	DB 6;	Length 18524;
Best Local Similarity		86.6%;	Pred. No. 1.2e-146;		
Matches		855;	Conservative 0;	Mismatches 132;	Indels 0; Gaps 0;
Qy	1	AAAGATGGGAAGAGCTGAGCTCCAGCTTGAAGTGCATGTAGAGGCCAAGAGCTGCAGA	60		
Db	981	AAGGATGGGAGAGAGCTGAGCTCCAGCTCGAAGGTGTGCATCGAGGCCACAGGCTGCAG	1040		
Qy	61	CGAGGCTGGTGGTGCAGCAGGCAAGAGCGGATGCCGGGGAATTACAGCTGCGAGGCC	120		
Db	1041	CGCAGGCTGGTGGTGCAGCAGGCAAGAGCGGATGCCGGGGAATTACAGCTGCGAGGCT	1100		
Qy	121	AGGGGCCAGAGGCTCTCTTCGCTGCACATCACAGAGCCCAAGATGATGTTCCAAAG	180		
Db	1101	GGGGGCCAGAGGCTCTCTTCGCTGCACATCACAGAGCCCAAGATGATGTTGCCAAG	1160		
Qy	181	GAGCAGTCAGTGCAATAGAGGTGCAGGCTGAGGGGGGGCCAGTGCCTATGCTGAGCTGT	240		
Db	1161	GACCAGTGGCACAACAGTCAGGTGCAGGCTGAGGCGAGGGGCCAGTGCCTGAGCTGC	1220		
Qy	241	GAGTGGCCAGGCCCCAGACGAGAGGTGACGTGGTACAAAGGATGGGAAGCTGAGCTCC	300		
Db	1221	GAGTGGCCAGGCCCCAGACGAGAGGTGATGTTGTACAAAGATGGGAAGAGCTGAGCTCC	1280		
Qy	301	AGCTCAAAAGTGGGATGAGGTCACAAAGGGTGCACACGGAGGCTGGTGTGCCACAGGG	360		
Db	1281	AGCTTGAAGTGCATGTAGAGGCCAAGGGCTGCAGACGAGAGGCTGGTGTGCAGCAGGA	1340		
Qy	361	GGCAAGCAGATGCTGGGGAGTACAGCTGTGAGGCTGGGGGCCAGAGAGTCTCTTTCCAC	420		
Db	1341	GGCAAGCAGGATGCTGGGGAGTACAGCTGTGAGGCTGGGGGCCAGAGAGTCTCTTTCCG	1400		
Qy	421	CTGCACATCACAGAGCCCAAGGGGTGTTTGCAGAGGAGCAGTCAAGTGCATTAAGGTG	480		
Db	1401	CTGCACATCACAGAGCCCAAGAGTATGTTTGCAGAGGAGCAGTCAAGTGCATTAAGGTG	1460		
Qy	481	CAGGCTGAGGCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCAGGCCAGAGGAG	540		
Db	1461	CAGGCTGAGGCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCAGGCCAGAGGAG	1520		
Qy	541	GTGACGTGTATCAGAGACGGGAAGAGCTGAGCTCCAGCTCAAAAGTACGATGAGGTC	600		
Db	1521	GTGACGTGTATCAGAGATGGGAAGAGCTGAGCTCCAGCTCAAAAGTGGGCATGAGGTC	1580		
Qy	601	AAGGGCTGCACACGAAGGCTGTGTAGTCAGCAGAGTGGGCAAGCAGAGTCTGGGGAGTAC	660		
Db	1581	AAAGGGTGCACACGAGAGCTGTGTGCTGCCACAGGCGGGCAAGCAGAGTCTGGGGAGTAC	1640		
Qy	661	AGCTGCGAGGCTGGGGGCCAGAGAGTCTCTTTTCAACTGCACATCACAGAGCCCAAGGCA	720		
Db	1641	AGCTGTGAGGCTGGGGGCCAGAGAGTCTCTTTTCAACTGCACATCACAGAGCCCAAGGG	1700		
Qy	721	GTGCTTTGCCAAGAGCAGTGTGTGCAATATGAGGTGCGGACTGAGGCGAGGGCCAGTGGC	780		
Db	1701	GTGCTTTGCCAAGAGCAGTCAAGTGTGCAATATGAGGTGCGGCTGAGGCGAGGGACCACTGGC	1760		
Qy	781	ACACTGAGCTGTGAGGTGGCCCGCCAGAGAGGTGACGTGGTGAACAAGGATGGGAAG	840		
Db	1761	ATGCTGAGCTGTGAGGTGGCCCGCCAGAGCCCGCCAGACAGAGGTGACGTGGTGAACAAGCAGGAAG	1820		
Qy	841	AAGCTGAGCTTCAGTTGGAAGTGGCAGATAGAGGCTGGGGTGCATGCGGCGAGCTGGTG	900		
Db	1821	AAGCTGAGCTTCCAGCTCAAAAGTACGCAATGGAGGTCAAGGGGCTGCACACGAAGGCTGGTA	1880		
Qy	901	GTGCAGCAGGCGCCAGGACAGATGCTGGGGAGTACACCTGTGAGGCTGGGGGCCAGCGG	960		
Db	1881	GTGCAGCAGGTTGGGCAAGCAGATGCTGGGGAGTACAGCTGCGAGGCTGGGGGCCAGAGA	1940		
Qy	961	CTCTCTTCCACTGGATGTTTCAGAG	987		
Db	1941	GTCTCTTTCAACTGCATCACAGAG	1967		


```
RESULT 3
HSA314900      19475 bp   DNA      linear      PRI 21-MAY-2002
LOCUS          Homo sapiens partial OBSCN gene for obscurin, exons all-a16.
DEFINITION
ACCESSION      AJ314900
VERSION        AJ314900.1  GI:21104329
KEYWORDS       OBSCN gene; obscurin.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS        Young, P., Ehler, E. and Gautel, M.
TITLE          Obscurin, a giant sarcomeric Rho-GEF protein involved in sarcomere
               assembly
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 19475)
AUTHORS        Gautel, M.S.
TITLE          Direct Submission
JOURNAL        Submitted (22-MAY-2001) Gautel M.S., Physikalische Biochemie,
               Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse
               11, Dortmund, 44227, GERMANY
FEATURES       Location/Qualifiers
               1..19475
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               4472..18215
               /gene="OBSCN"
               join(<4472..4747,585..5860,6702..6977,7817..8092,
               11254..11529,17940..>18215)
               /gene="OBSCN"
               join(<4472..4747,585..5860,6702..6977,7817..8092,
               11254..11529,17940..>18215)
               /gene="OBSCN"
               /codon_start=3
               /product="obscurin"
               /protein_id="CAC85747.1"
               /db_xref="GI:21104330"
               /db_xref="UniProt/TREMBL:O8NH7"
               /translation="PKVFAKDVVAHSEVQAEASATLSCVEVAQAQTEVMWYKDGKK
               LSSLVKVAEAGKRLVVOQAQKTDAGDSCEARGVRSFRLHTEPRKMFQKQS
               VHNVEQAEASAKMLSCVEVAQAQTEVTWYKDKLSSSKVMEVKGCTRLVLVPOAG
               KADAGEYCEAGGQVRSFHLHTEPKGVFAKEQSVHNEVQAEAGTTLSCVEVAQPT
               EWTYKDGKLLSSSKVMEVKGCTRLVVOQVQKADAGEYSCGAGQVRSFQHLITE
               PKAVFAKQLVHNEVTRFAGASATLSCVEVAQAQTEVTWYKDKLSSSKVRIEAGC
               MROLVVOQAQADAGEYTCGAGGQRLSPHLVSEPKAVFAKEQLAHRKQVQAEALAT
               LSCVEVAQAQTEVTWYKDKLSSSKVMEVAGCTRLVVOQAQADAGEYSCGAGQ
               LRSFLDVAEPKVFQAEQPVHREVQQAQAGATTLSCVEVAQAQTEVMWYKDGKLLSFS
               SKVRMEAAGCTRLVVOQAQVAGVSCGAGSQRSLSFHLHVA"
               <4472..4747
               /gene="OBSCN"
               /note="a11"
               4748..5584
               /gene="OBSCN"
               /note="a11"
               5585..5860
               /gene="OBSCN"
               /note="a12"
               5861..6701
               /gene="OBSCN"
               /note="a12"
               6702..6977
               /gene="OBSCN"
               /note="a13"
               6978..7816
               /gene="OBSCN"
               /note="a13"
               7817..8092
               /gene="OBSCN"
               /note="a14"
               8093..11253
               /gene="OBSCN"
               /note="a14"
               11254..11529
               /gene="OBSCN"
               /note="a15"
               11530..17939
               /gene="OBSCN"
               /note="a15"
               17940..>18215
               /gene="OBSCN"
               /note="a16"

exon
intron
exon
ORIGIN
Query Match      27,9%; Score 279.4; DB 9; Length 19475;
Best Local Similarity 99.6%; Pred. No. 1.8e-46;
Matches 280; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 153 CACAGAGCCCAAGATGATGTTTGC AAAAGGAGCAGTCAGTCATATGAGGTGCAGGCTGA 212
Db      |||||
5580 CTCAGAGCCCAAGATGATGTTTGC AAAAGGAGCAGTCAGTCATATGAGGTGCAGGCTGA 5639
QY 213 GGC GGGGGCCAGTCCCATGCTGAGCTGTGAGGTGGCCAGCCCGAGCGAGGTGACGTG 272
Db      |||||
5640 GGC GGGGGCCAGTCCCATGCTGAGCTGTGAGGTGGCCAGCCCGAGCGAGGTGACGTG 5699
QY 273 GTACAAGGATGGGGAAGAAGCTGAGCTCCAGCTCAAAAGTGGGCATGAGGTCAAAAGGCTG 332
Db      |||||
5700 GTACAAGGATGGGGAAGAAGCTGAGCTCCAGCTCAAAAGTGGGCATGAGGTCAAAAGGCTG 5759
QY 333 CACACGAGGCTGCTGCTGCCACAGCGGGGCAAGCAGATGCTGGGAGTACACTGTGA 392
Db      |||||
5760 CACACGAGGCTGCTGCTGCCACAGCGGGGCAAGCAGATGCTGGGAGTACACTGTGA 5819
QY 393 GGCTGGGGGGCCAGAGAGTCTCTTCCACCTGCACATCACAG 433
Db      |||||
5820 GGCTGGGGGGCCAGAGAGTCTCTTCCACCTGCACATCACAG 5860

RESULT 4
LOCUS          AL353593
DEFINITION     Human DNA sequence from clone RP5-1139B12 on chromosome 1q42.1-43,
               complete sequence.
ACCESSION      AL353593
VERSION        AL353593.33  GI:18673899
KEYWORDS       HTG.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 135964)
AUTHORS        Almeida, J.
TITLE          Direct Submission
JOURNAL        Submitted (13-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
               Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
               humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
               On Feb 14, 2002 this sequence version replaced gi:17977879.
               During sequence assembly data is compared from overlapping clones.
               Where differences are found these are annotated as variations
               together with a note of the overlapping clone name. Note that the
               variation annotation may not be found in the sequence submission
               corresponding to the overlapping clone, as we submit sequences with
               only a small overlap as described above.
               This sequence was finished as follows unless otherwise noted: all
               regions were either double-stranded or sequenced with an alternate
               chemistry or covered by high quality data (i.e., phred quality >=
               30); an attempt was made to resolve all sequencing problems, such
               as compressions and repeats; all regions were covered by at least
               one plasmid subclone or more than one M13 subclone; and the
               assembly was confirmed by restriction digest. The following
               abbreviations are used to associate primary accession numbers given
               in the feature table with their source databases: Em:, EMBL; Sw:,
               SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; information on the WORMPEP
```


database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>
RP5-1139B12 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>
VECTOR: pCVPAC2
This sequence is the entire insert of clone RP5-1139B12.

[illegible]

Best Local Similarity 57.4%; Pred. No. 4.4e-40;					
Matches 510; Conservative					
					Pred. No. 4.4e-40; Matches 510; Conservative
					Mismatches 366; Indels 12; Gaps 3;
Qy	75	GCACGAGCAGGCAAGACGGATCCCGGGACTACAGCTGCGAGGCCAAGGGCCAGAGGGT	134		
Dd	1874	GGAAGTGCCCGCACGCTGCACGGGGCGCAGCTGAAGTTCTGTGGCCACAAGCATGTGAG	1933		
Qy	135	CTCCTTCGCGCTGCACATCACAGAGCCCAGAATGATGTTTTGCAAAGAGCAGTCAGTGCA	194		
Dd	1934	CAGCATCCGGATGGAGTCCGGCGGCCCCAGGGCTGACTGCCAAACAAGCCGACGCGC	1993		
Qy	195	TAAAT-----GAGGTGCAAGCTAGCGGGGGGCCAGTGCATCTGAGCTGTGAGGTGGC	248		
Dd	1994	AGCTGCCCGGGAGGTCTGCTCTGGCTGTCACGAGGAGGCGCAGCTCTCTGCTGAGCTGT	2053		
Qy	249	CCAGGCCACAGACGGAGTGAAGTGTGTAACAAGATGGGAAGAGCTCAGCTCCAGCTCAA	308		
Dd	2054	AGATCAGGCTGCGGCTGTGACGTGGCTGAAGGATGTGCGACACTGTCCCCAGGCCCAA	2113		
Qy	309	AGTGGGCATGGAGTCAAAAGGTGTGCAACGGAGGCTGGTGTCTGCCACAGCGCGCAAGC	368		
Dd	2114	GTATGAGTGCAGGCATCGCGCGGGCGGGTGTCTCTGTGCGAGATGTGCGCCGGGA	2173		
Qy	369	AGATGCTGGGGAGTACAGCTGTGAGGCTGGGGCCAGAGAGTCTCTTCCACCTGCACAT	428		
Dd	2174	CSATGCAGGCGCTCTACGAGTGCCTCAGCGCGCGGGGCCGATCCGCTACCAGCTCTCCGT	2233		
Qy	429	CACAGAGCCCAAGGGGTGTTTCGGAAGGAGCAGTCAAGTGCATAATGAGGTGCAGGCTGA	488		
Dd	2234	GCAAGGCTCGCGGCTTCTTCGCAAGAGNACTGGCGGCGAGCTGT---GTGATGCCGT	2290		
Qy	489	GGCGGGACCACTGCCATGCTGAGCTGTGAGTGGGCCACGCCCCAGACGAGGTGACGCTG	548		
Dd	2291	GGCTGGGGGCCCGCGCAGTTTTCAGTGTGAGACCTCCGAAGCCCACTGCCACGTGCACTG	2350		
Qy	549	GTACAAGACGGGAAGAGCTGAGCTCCAGCTCAAAGTACGATGCGAGGTCAAGGGCTG	608		
Dd	2351	GTACAAGGATGGCATGGAGCTGGGGCCACTCCGGTGAAGCGCTTCTTGACGAGGATGTGG	2410		
Qy	609	CACACG---AAGGCTGTAGTGACAGAGTGGGCAAGACAGATGCTGGGAGGTACAGCTG	665		
Dd	2411	GACCGGACCGGCTGTGTGCAGCCACAGTTCACGAGGAGGATGAAGCNACTTACTCTTG	2470		
Qy	666	CGAGCTGGGGGCCAGAGAGTCTCTTTCACTGSCATCACAGAGCCCCAAGCAGTGT	725		
Dd	2471	CCCGTGGGCGAGGACTCTGTGACTTCCGGCTCCCGCTCTCTGAGCCCCAAGCGGTGT	2530		
Qy	726	TGCCAAGGAGCATTGGTGCATAATGAGTGCAGACTGAGGCAAGGGCCAGTCCACACT	785		
Dd	2531	TGCCAAGGAGCACCAGCGGTGACGGAGGTGTCAGGCTGAGGTGGGGGCCAGCGCCAGCT	2590		
Qy	786	GAGCTGTGAGTGGCGCCAGGCCACAGAGGTGACGTGTGTACAAGATGGGAAGAAGCT	845		
Dd	2591	GAGCTGTGAGTGGCGCCAGGACACAGATGAGGTGACGTGTGTACAAGAACGGGAAGTT	2650		
Qy	846	GAGCTCCAGTTCCAAAAGTGCACATAGAGGCTGCGGGCTGCATGCGGCGAGCTGGTGGTGA	905		
Dd	2651	GAGCTCCAGTCCAAAGTGCACGTGGAGGCTGTGGGCTGTATGCGGAGGCTGGTGGTGA	2710		
Qy	906	GCAGGAGGCCAGGCAAGTCTGGGAGGTACACTGTGAGGCTGGGGG 953			
Dd	2711	GCAGGTGGGCCAGGCAGACTCCGGATAGTACAGCTGTGAAGCCAGGTG 2758			
RESULT 6					
BD160445					
LOCUS					
DEFINITION Primer for synthesizing full-length cDNA and use thereof.					
ACCESSION BD160445					
VERSION BD160445.1 GI:27866203					
KEYWORDS JP 2002191363-A/15286.					
SOURCE Homo sapiens (human)					
ORGANISM Homo sapiens					


```
misc_feature 25690..34815
              /note="assembly_name:Contig10"
misc_feature 34916..45742
              /note="assembly_name:Contig11"
misc_feature 45843..62987
              /note="assembly_name:Contig12
clone_end:SP6
vector_side:right"
misc_feature 63088..87584
              /note="assembly_name:Contig13"
misc_feature 87685..114765
              /note="assembly_name:Contig14"
misc_feature 114866..142334
              /note="assembly_name:Contig15"
misc_feature 142435..174612
              /note="assembly_name:Contig16"

ORIGIN
Query Match      21.8%; Score 218.6; DB 2; Length 174612;
Best Local Similarity 86.1%; Pred. No. 2.3e-34;
Matches 242; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 705 CACAGAGCCCAAGGCAGTGTTCACAGGAGCAGTTGGTCATATGAGGTGCGGACTGA 764
DB 95992 CCCAGAGCCCAAGGTGGTGTTCACAGGAGCAGCAGCACACAGGAGGTGCAGGCTGA 96051
QY 765 GCGAGGGGCGAGTGCACACTAGCTGTGAGTGGCCAGGCCAGCAGACAGAGGTGACGTG 824
DB 96052 GCGGGGGGCCAGTGCACGCTGAGCTGGAGGTGGCCAGGCCAGCAGACAGAGGTGACGTG 96111
QY 825 GTACAAGAGTGGGAAGAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGGCTCGGGGCTG 884
DB 96112 GTACAAGAGTGGGAAGAGCTGAGTTCAGCTCGAAAGTGGCGTGGAGGCGGTGGGCTG 96171
QY 885 CATGGCGGAGCTGTGTGTGCAGCAGGCGAGGCCAGCAGATGCTGGCGAGTACACCTGTGA 944
DB 96172 CACAGGAGGCTGTGTGTGCAGCAGGCGGCGCAGCAGGCGGGGAGTACAGTGTGGA 96231
QY 945 GCGTGGGGCGCAGCGGCTCTCTCCACCTGGAGTGTTCAG 985
DB 96232 GCGAGGGGTGAGCAGCTCTCTCCGCTGCGAGTGGCGAG 96272

RESULT 11
LOCUS CQ851220/c 3956 bp DNA linear PAT 23-AUG-2004
DEFINITION Sequence 1689 from Patent EP1447413.
ACCESSION CQ851220
VERSION CQ851220.1 GI:51509432
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isegai,T., Yamamoto,J., Nishikawa,T., Isono,Y., Sugiyama,T.,
Otsuki,T., Wakamatsu,A., Ishii,S., Nagai,K. and Irie,R.
TITLE Full-length human cDNA
JOURNAL Patent: EP 1447413-A 1689 18-AUG-2004;
RESEARCH Association for Biotechnology (JP)
FEATURES
          Location/Qualifiers
          1..3956
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"

ORIGIN
Query Match      21.8%; Score 218; DB 6; Length 3956;
Best Local Similarity 84.5%; Pred. NO. 5.7e-34;
Matches 245; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 144 CCTGCATCAGAGCCCAAGATCATCTTTGCAAGGAGCAGTCAGTGCATATGAGGT 203
```

```
DB 1141 CTTGTCCATCCAGAGCCCAAGGTGGTGTTCACAGGAGCAGCGCGCATGCAGGGAGGT 1082
QY 204 GCAGGCTGAGCGCGGGGCCAGTGCATGCTGAGCTGTGAGTGGCCAGGCCAGACGGA 263
DB 1081 GCAGCGGAGCGCGGGGCCAGTGCCTGAGCTGTGAGTGGCCAGGGGCGAGATGA 1022
QY 264 GGTGACCTGGTACAGAGATGGAGAGCTCAGCTCCAGCTCAAAAGTGGGCATGGAGGT 323
DB 1021 GGTGACATGGTACAGAGACGGGAGAGCTGAGCTCCAGCTCAAAAGTGGCGATGGAGGC 962
QY 324 CAAAGGGTGCACACGAGAGCTGGTCTGCCACAGCGGGGCGGCAAGCAGATGCTGGCGAGTA 383
DB 961 CAGCGGCTACACACGAGGCTGTGTAGTGCACAGCGGGGCGGCGGATGCTGGGAGTA 902
QY 384 CAGCTGTGAGCTGGGGGCCAGAGAGTCTCTTCACCTGCACATCACAG 433
DB 901 CAGCTGTGAGCGGGGGCCCAACGGCTGTCTTCGCTGCGACGTGGCGAG 852

RESULT 12
LOCUS AK128447/c 3956 bp mRNA linear PRI 19-FEB-2004
DEFINITION Homo sapiens cDNA FLJ46590 fis, clone THYMU304441.
ACCESSION AK128447
VERSION AK128447.1 GI:34535823
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Ninomiya,K., Wagatsuma,M., Kanda,K., Kondo,H., Yokoi,T.,
Kodaira,H., Furiya,T., Takahashi,M., Kikkawa,E., Omura,Y., Abe,K.,
Kamihara,K., Katsuta,N., Sato,K., Tanikawa,M., Yamazaki,M.,
Sugiyama,T., Irie,R., Otsuki,T., Sato,H., Ota,T., Wakamatsu,A.,
Ishii,S., Yamamoto,J., Isono,Y., Kawai-Hio,Y., Saito,K.,
Nishikawa,T., Kimura,K., Yamashita,H., Matsuo,K., Nakamura,Y.,
Sekine,M., Kikuchi,H., Murakawa,K., Kanehori,K.,
Takahashi-Fujii,A., Oshima,A., Sugiyama,A., Kawakami,B., Suzuki,Y.,
Sugano,S., Nagahari,K., Masuho,Y., Nagai,K. and Isegai,T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3956)
AUTHORS Isegai,T. and Yamamoto,J.
JOURNAL Direct Submission
TITLE Submitted (15-JUL-2003) Takao Isegai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail:genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: Reverse Proteomics Research Institute, HRI and
RAB.
FEATURES
          Location/Qualifiers
          1..3956
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="THYMU304441"
            /tissue_type="thymus"
            /clone_lib="THYMU3"
            /note="Cloning vector: pME18SFL3"
            1014..1415
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="BAC87445.1"
              /db_xref="GI:34535824"
              /translation="MSPSPGPGPPHSSAHWPPPPPPAPPAPAPWOTPPWALGHTRR
              DAGSEAPEQTQPMQCQSDSQGLSIIKQCTVTHRAPPTTRAKLRLHVGSTIIMSLG"

CDS
            1014..1415
              /note="unnamed protein product"
              /codon_start=1
              /protein_id="BAC87445.1"
              /db_xref="GI:34535824"
              /translation="MSPSPGPGPPHSSAHWPPPPPPAPPAPAPWOTPPWALGHTRR
              DAGSEAPEQTQPMQCQSDSQGLSIIKQCTVTHRAPPTTRAKLRLHVGSTIIMSLG"
```


ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE
AUTHORS Young, P., Ehler, E. and Gautel, M.
TITLE Obscurin, a giant sarcomeric rho-GEF protein involved in sarcomere assembly
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 10091)
AUTHORS Gautel, M.S.
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2001) Gautel M.S., Physiokalische Biochemie,
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse
11, Dortmund, 44227, GERMANY
FEATURES
source
1..10091
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
1265..6155
/gene="OBSCN"
/join<1265..1540,5880..>6155)
/gene="OBSCN"
/join<1265..1540,5880..>6155)
/gene="OBSCN"
/codon_start=3
/product="Obscurin"
/protein_id="CAC85749.1"
/db_xref="GI:21104332"
/db_xref="UniProt/TREMBL:Q8NH96"
/translation="PKAVFAKEQPAHREVQAEAGASATLSCVEAAQOTETVWYKDGKK
LSSSIKVHVEAGCTRLVLVQAGQADTGEYSCAGGQQLSFRLOVAFPKAVFAKEQP
ASREVQAGAGTSATLSCVEAAQOTETVWYKDGKLSKSSKVRMEAVGCTRLVLVQAEAG
QADGEYSCAGQDRLSPLHVA"
1265..1540
/gene="OBSCN"
/note="B"
1541..5879
/gene="OBSCN"
/number=16
5880..>6155
/gene="OBSCN"
/number=17
ORIGIN
Query Match 21.5%; Score 215.4; DB 9; Length 10091;
Best Local Similarity 84.1%; Pred. No. 1.6e-33;
Matches 243; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 697 CTGCACATCACAGAGCCAGCAGTGTTCGCAAGGAGCAGTGTGTCATATAGGTG 756
Db 1252 CTGCTTCCCCAGAGCCCAAGCGGTGTTTGTAAAGGAGCAGCCACACAGGAGGTG 1311
QY 757 CGGACTGAGGAGGGGCGGAGTGCACACTGAGCTGTGAGTGGCCGAGCCAGCCAGACAGAG 816
Db 1312 CAGGCTGAAGCAGGGGGCTAGCGCCACACTGAGCTGTGAGTGGCCGAGCCAGACAGAG 1371
QY 817 GTGACGTGGTACAGGATGGGAAGAGCTGAGCTCCAGTTGCAAAAGTCGCAATAGAGGCT 876
Db 1372 GTGACTTGGTACAGGACGGGAAGAGCTGAGCTCCAGTTGCAAAAGTCGATGGAGGCT 1431
QY 877 GCGGCTGATCGGCGAGCTGTGTGTCAGCAGGAGGAGCCAGGAGAGTGTGGGAGGTAC 936
Db 1432 GCAGCTGTACACGGAGGCTGTGTGTGTCAGCAGGAGGAGCCAGGAGACACCGGAGAGTAT 1491
QY 937 ACCTGTGAGGCTGGGGGCGAGGGCTCTCTTCCACCTGGATGTTTCAG 985
Db 1492 AGCTGTGAGGCTGGGGGCGAGGAGCTCTCTTCCGCTGCAAGTGGCAG 1540
RESULT 15
AC026657
LOCUS AC026657 164766 bp DNA linear HTG 01-SEP-2000

DEFINITION Homo sapiens chromosome 1 clone RP11-245P10, WORKING DRAFT
SEQUENCE, 31 unordered pieces.
ACCESSION AC026657
VERSION AC026657.4 GI:9958202
KEYWORDS HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 164766)
REFERENCE
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 164766)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT
On Sep 1, 2000 this sequence version replaced gi:7637349.
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0245P10
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-terminator ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 139884 bases at least Q40
Consensus quality: 147686 bases at least Q30
Consensus quality: 151469 bases at least Q20
Insert size: 163000; agarose-efp
Insert size: 161074; sum-of-contigs
Quality coverage: 3.60 in Q20 bases; agarose-efp
Quality coverage: 3.92 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1124: contig of 1124 bp in length
* 1125 1224: gap of unknown length
* 1225 1225: contig of 1808 bp in length
* 3033 3132: gap of unknown length
* 3133 4493: contig of 1361 bp in length
* 4494 4594: gap of unknown length
* 4594 5860: contig of 1267 bp in length
* 5861 5961: gap of unknown length
* 5961 7670: contig of 1710 bp in length
* 7671 7771: gap of unknown length
* 7771 9669: contig of 1899 bp in length
* 9670 9769: gap of unknown length
* 9770 11715: contig of 1946 bp in length
* 11716 11815: gap of unknown length
* 11816 14243: contig of 2428 bp in length
* 14244 14343: gap of unknown length
* 14344 16887: contig of 2544 bp in length
* 16888 19287: gap of unknown length
* 19288 19247: contig of 2260 bp in length
* 19248 21375: contig of 2028 bp in length
* 21376 21475: gap of unknown length
* 21476 25025: contig of 3550 bp in length
* 25026 25125: gap of unknown length


```
* 25126 28051: contig of 2926 bp in length
* 28052 28151: gap of unknown length
* 28152 32054: contig of 3903 bp in length
* 32055 32154: gap of unknown length
* 32155 35716: contig of 3562 bp in length
* 35717 35816: gap of unknown length
* 35817 40891: contig of 5075 bp in length
* 40892 40991: gap of unknown length
* 40992 44027: contig of 3036 bp in length
* 44028 44127: gap of unknown length
* 44128 48899: contig of 4772 bp in length
* 48900 54255: contig of 5256 bp in length
* 54256 54356: gap of unknown length
* 54357 58995: contig of 4639 bp in length
* 58996 59094: gap of unknown length
* 59095 64390: contig of 5296 bp in length
* 64391 70865: contig of 6375 bp in length
* 70866 78667: contig of 7702 bp in length
* 78668 87207: contig of 8440 bp in length
* 87208 87307: gap of unknown length
* 87308 97859: contig of 10551 bp in length
* 97860 109520: contig of 11562 bp in length
* 109521 120908: gap of unknown length
* 120909 121009: contig of 11288 bp in length
* 121010 141477: contig of 20469 bp in length
* 141478 162672: contig of 21095 bp in length
* 162673 162773: gap of unknown length
* 162774 163975: contig of 1202 bp in length
* 163976 164075: gap of unknown length
* 164076 164766: contig of 692 bp in length.
FEATURES
    source
        1..164766
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="1"
            /clone="RP11-245P10"
        1..1124
            /note="assembly_name:Contig14"
        1225..3032
            /note="assembly_name:Contig21"
        3133..4493
            /note="assembly_name:Contig23"
        4594..5860
            /note="assembly_name:Contig24"
        5961..7670
            /note="assembly_name:Contig25"
        7771..9669
            /note="assembly_name:Contig26"
        9770..11715
            /note="assembly_name:Contig27"
        11816..14243
            /note="assembly_name:Contig28"
        14344..16887
            /note="assembly_name:Contig29"
        16988..19247
            /note="assembly_name:Contig30"
        19348..21375
            /note="assembly_name:Contig31"
        21476..25025
            /note="assembly_name:Contig32"
        25126..28051
            /note="assembly_name:Contig33"
        28152..32054
            /note="assembly_name:Contig34"
        32155..35716
```

```
misc_feature /note="assembly_name:Contig35" 35817..40891
misc_feature /note="assembly_name:Contig36" 40992..44027
misc_feature /note="assembly_name:Contig37" 44128..48899
misc_feature /note="assembly_name:Contig38" 49000..54255
misc_feature /note="assembly_name:Contig39" 54356..58994
misc_feature /note="assembly_name:Contig40" 59095..64390
misc_feature /note="assembly_name:Contig41" 64491..70865
misc_feature /note="assembly_name:Contig42" 70966..78667
misc_feature /note="assembly_name:Contig43" 78768..87207
misc_feature /note="assembly_name:Contig44" 87308..97859
misc_feature /note="assembly_name:Contig45" 97959..109520
misc_feature /note="assembly_name:Contig46" 109621..120908
misc_feature /note="assembly_name:Contig47" 121009..141477
misc_feature /note="assembly_name:Contig48" 141578..162672
misc_feature /note="assembly_name:Contig49" 162773..163974
misc_feature /note="assembly_name:Contig9" 164075..164766
misc_feature /note="assembly_name:Contig7"

ORIGIN
Query Match      20.7%; Score 207.6; DB 2; Length 164766;
Best Local Similarity 84.5%; Pred. No. 3.9e-32;
Matches 245; Conservative 0; Mismatches 44; Indels 1; Gaps 1;

Qy 144 CCTGCATCACAGAGCCCAAGATGATGTTTGCNAGGAGCAGTCAGTGCATATAGGT 203
Db 56107 CTTGTCCATCCAGAGCCCAAGGTGGTGTTCNAGGAGCAGCCGCATGCAGGGAGGT 56166

Qy 204 GCAGGCTGAGCGGGGCCAGTGCCATGCTGAGCTGAGGTGGCCAGCCAGACCGGA 263
Db 56167 GCAGGCGAGCGGGGCCAGTGCCAGCTGAGCTGAGGTGGCCAGCCAGATGGA 56226

Qy 264 GGTGACGTGTACAGGATGGGAGAGCTGAGCTCCAGTCAAAAGTGGGCATGGAGGT 323
Db 56227 GGTGACATGGTACAAAGACCGGAGAGCTGAGCTCCAGCTCAAAAGTGGGCATGGAGGC 56286

Qy 324 CAAAGGCTGCACACGAGGCTGGTGTCCACAGGGGGCCNAGCAGATGCTGGGAGTA 383
Db 56287 CAGCGGCTACACACGAGGCTGGTAGTGACAGAGGGGGCCAGCGGATGCTGGGAGTA 56346

Qy 384 CAGCTGTAGGCTGGGGGCCAGAGAGTCTCTTCCACTGCATCACAG 433
Db 56347 CAGCTGTAGGC-GGGGGCCACGCGTGTCTTCCGCTGCACGTGCAG 56395
```

Search completed: March 21, 2005, 05:34:31
Job time : 4521.62 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:32:32 ; Search time 627.052 Seconds
(without alignments)
9503.485 Million cell updates/sec

Title: US-10-077-130-4_COPY_22500_23500

Perfect score: 1001

Sequence: 1 ggcagtcagtgacactggc.....ctgtcaaggtcgtggacct 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	3225	17 US-10-311-034-46	Sequence 46, Appl
2	1001	100.0	4175	15 US-10-307-019-2	Sequence 2, Appli
3	1001	100.0	4936	17 US-10-182-243-14	Sequence 14, Appl
4	1001	100.0	5454	17 US-10-415-011-44	Sequence 44, Appl
5	1001	100.0	7893	13 US-10-077-130-3	Sequence 3, Appli
6	1001	100.0	8106	13 US-10-077-130-1	Sequence 1, Appli
7	1001	100.0	23907	13 US-10-077-130-6	Sequence 6, Appli
8	1001	100.0	24120	13 US-10-077-130-6	Sequence 4, Appli
9	999.4	99.8	5007	15 US-10-307-019-3	Sequence 3, Appli
10	999.4	99.8	5207	9 US-09-858-664A-1	Sequence 1, Appli
11	999.4	99.8	5207	14 US-10-274-978-1	Sequence 1, Appli

12	999.4	99.8	5207	14	US-10-274-978-3	Sequence 3, Appli
13	999.4	99.8	5207	17	US-10-697-263-1	Sequence 1, Appli
14	999.4	99.8	5207	17	US-10-697-263-3	Sequence 3, Appli
15	999.4	99.8	5207	18	US-10-921-168-1	Sequence 1, Appli
16	999.4	99.8	5207	18	US-10-921-168-3	Sequence 3, Appli
17	999.4	99.8	7928	15	US-10-307-019-5	Sequence 5, Appli
18	696	69.5	1253	17	US-10-425-114-26278	Sequence 26278, A
19	151.6	15.1	9930	17	US-10-210-130-33	Sequence 33, Appli
20	148.4	14.8	1627	17	US-10-425-114-26265	Sequence 26265, A
21	148.4	14.8	7789	17	US-10-333-314-38	Sequence 38, Appli
22	146.8	14.7	9698	17	US-10-210-130-37	Sequence 37, Appli
23	146.8	14.7	10122	17	US-10-210-130-35	Sequence 35, Appli
24	140.4	14.0	9807	17	US-10-379-381-1	Sequence 1, Appli
c 25	131	13.1	458	13	US-10-027-632-272062	Sequence 272062, A
c 26	131	13.1	458	17	US-10-027-632-272062	Sequence 272062, A
c 27	129.8	13.0	428	13	US-10-027-632-272063	Sequence 272063, A
c 28	129.8	13.0	428	17	US-10-027-632-272063	Sequence 272063, A
29	105.6	10.5	62805	17	US-10-379-381-3	Sequence 3, Appli
30	97.8	9.8	860	17	US-10-210-130-39	Sequence 39, Appli
31	92	9.2	2224	14	US-10-059-585-11	Sequence 11, Appli
32	92	9.2	2226	18	US-10-335-053-219	Sequence 219, Appl
33	90.4	9.0	2105	18	US-10-468-406-17	Sequence 17, Appli
34	90.4	9.0	2132	17	US-10-262-445-118	Sequence 118, Appl
c 35	89.8	9.0	3133	18	US-10-437-963-41514	Sequence 41514, A
36	86.2	8.6	1452	17	US-10-425-114-26649	Sequence 26649, A
37	86.2	8.6	1581	18	US-10-425-115-182879	Sequence 182879, A
38	85	8.5	1984	17	US-10-425-114-19039	Sequence 19039, A
39	85	8.5	2062	17	US-10-425-114-34609	Sequence 34609, A
40	85	8.5	2233	18	US-10-425-115-179544	Sequence 179544, A
41	84.2	8.4	1584	17	US-10-302-172-205	Sequence 205, Appl
42	82.8	8.3	1282	10	US-09-769-970-12	Sequence 12, Appli
43	82.8	8.3	1282	17	US-10-305-720-953	Sequence 953, Appl
44	82.8	8.3	1282	18	US-10-656-598-5	Sequence 5, Appli
45	82.8	8.3	1282	18	US-10-643-795A-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1
US-10-311-034-46
; Sequence 46, Application US/10311034
; Publication No. US20040023242A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: YUE, Henry
; APPLICANT: LAL, Preeti
; APPLICANT: BANDMAN, Olga
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Yan
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: YAO, Monique G.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: GREENWALD, Sara R.
; APPLICANT: RAMKUMAR, Javalaxmi
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: KEARNEY, Liam
; APPLICANT: BURFORD, Neil
; APPLICANT: NGUYEN, Danniell B.
; APPLICANT: TANG, Y. Tom
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: HE, Ann
; APPLICANT: THORNTON, Michael
; APPLICANT: HAFALIA, April
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: LO, Terence P.
; APPLICANT: KHAH, Farrah A.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: AZIMZAI, Yalda

QY	661	GC	CTGGCCGCGCAAGATCATCCCTACCA	CCCAAGCAAGACAGCAGTGTGCGCGAA	720
Db	2323	GC	CTGGCGCGCAAGATCATCCCTACCA	CCCAAGCAAGACAGCAGTGTGCGCGAA	2382
QY	721	TAC	GAGGCCCTCAAGGGCCTGGCCACCG	CGACCTGGCCCGTGCACGAGCCTACCTC	780
Db	2383	TAC	GAGGCCCTCAAGGGCCTGGCCACCG	CGACCTGGCCCGTGCACGAGCCTACCTC	2442
QY	781	AG	CCCCCGGCACCTGTGTGCTCATCTT	GGAGCTGTGTCTTGGGCCCGAGCTGTCTCCCTGC	840
Db	2443	AG	CCCCCGGCACCTGTGTGCTCATCTT	GGAGCTGTGTCTTGGGCCCGAGCTGTCTCCCTGC	2502
QY	841	CT	GCCGAGAGGGCTCCTACTCAGAA	TCCGAGGTGAAGCACTACTGTGGCAGATGTTG	900
Db	2503	CT	GCCGAGAGGGCTCCTACTCAGAA	TCCGAGGTGAAGCACTACTGTGGCAGATGTTG	2562
QY	901	AG	TGCCACCCAGTACCTGCACAAC	CAGCAGCATCCTGCACCTGGACCTGAGGTC	960
Db	2563	AG	TGCCACCCAGTACCTGCACAAC	CAGCAGCATCCTGCACCTGGACCTGAGGTC	2622
QY	961	AT	GATCATCATCCGAATACAACTGT	CAAGTCTGTGAAGCT	1001
Db	2623	AT	GATCATCATCCGAATACAACTGT	CAAGTCTGTGAAGCT	2663

```

RESULT 2
US-10-307-019-2
; Sequence 2, Application US/10307019
; Publication No. US20030108533A1
; GENERAL INFORMATION:
; APPLICANT: zeng, wenlin
; APPLICANT: Stanton, Lawrence
; APPLICANT: SCIOS, INC.
; TITLE OF INVENTION: MEDIATORS OF SIGNAL TRANSDUCTION
; FILE REFERENCE: SCIOS.021DV1
; CURRENT APPLICATION NUMBER: US/10/307,019
; CURRENT FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/548,473
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: 60/129,552
; PRIOR FILING DATE: 1999-04-16
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4175
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(4053)
US-10-307-019-2

Query Match      100.0%; Score 1001; DB 15; Length 4175;
Best Local Similarity 100.0%; Pred. NO. 2.3e-261;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  GGCCAGTCAGTGACCTGCGCTGCCAGAGTGTCAGCCCGCAGCAGCTGCCAGGCCACCTGG 60
DB      2578  GGCCAGTCAGTGACACTGCGCTGCCAGAGTGTCAGCCCGCAGCAGCTGCCAGGCCACCTGG 2637

QY      61  AGCAAAAGCGGAGCCGCCCTTGGAGAGCAGCAGCCGGTGCTCATCTCTGCCACCTCAAG 120
DB      2638  AGCAAAAGCGGAGCCGCCCTTGGAGAGCAGCAGCCGGTGCTCATCTCTGCCACCTCAAG 2697

QY      121  AACTTCACGTTCTGACCATCTGTGGTGTTGGCTGAGGACCTGGGTGTGTACACCTGC 180
DB      2698  AACTTCACGTTCTGACCATCTGTGGTGTTGGCTGAGGACCTGGGTGTGTACACCTGC 2757

QY      181  AGCGTGAGCAATCCGCTGGGGAAGTGAACCAACCAACCGGCGTCTCTCGGAAGCAGAGCGC 240
DB      2758  AGCGTGAGCAATCGCTGGGGAAGTGAACCAACCAACCGGCGTCTCTCGGAAGCAGAGCGC 2817

```



```
OY 241 CCCTCATCTTCCCATGCCCCGATATCGGGAGGTGTACGGATGGGGTGTCTGTGTC 300
Db 2818 CCCTCATCTTCCCATGCCCCGATATCGGGAGGTGTACGGATGGGGTGTCTGTGTC 2877
OY 301 TGGAAAGCCGCGTGAATCTCTACGGCCCTGTGACCTACATTGTGCAAGTGTAGAGGC 360
Db 2878 TGGAAAGCCGCGTGAATCTCTACGGCCCTGTGACCTACATTGTGCAAGTGTAGAGGC 2937
OY 361 GGCAGCTGGACACACACTGGCCCTCCGACATCTTTTGACTGCTACTGACCAAGCAAGTTC 420
Db 2938 GGCAGCTGGACACACACTGGCCCTCCGACATCTTTTGACTGCTACTGACCAAGCAAGTTC 2997
OY 421 TCCCGGGGTGGACCTTACACCTTCGCGACGGCATGTGTGAGCAAGCAAGCAAGTTC 480
Db 2998 TCCCGGGGTGGACCTTACACCTTCGCGACGGCATGTGTGAGCAAGCAAGCAAGTTC 3057
OY 481 TACAGCAGCCCTCGAGCAAGTCTCTCGGAGGGCCGAGCCACTTGGCTCTGAGGAG 540
Db 3058 TACAGCAGCCCTCGAGCAAGTCTCTCGGAGGGCCGAGCCACTTGGCTCTGAGGAG 3117
OY 541 GAGAGCCAGGGCGGTGACGCCCAACCCCTGCCAGCAAAAGACCTTGGCATTCAGAGCA 600
Db 3118 GAGAGCCAGGGCGGTGACGCCCAACCCCTGCCAGCAAAAGACCTTGGCATTCAGAGCA 3177
OY 601 CAGATCCAGAGGGCGGCTTACAGCGTGTGCGGCAATGCTGGGAGAGAGCCAGCGGCG 660
Db 3178 CAGATCCAGAGGGCGGCTTACAGCGTGTGCGGCAATGCTGGGAGAGAGCCAGCGGCG 3237
OY 661 GCGTGGCGCCGCAAGATCATCCCTACCAACCCCAAGGCAAGACAGCAGTGTGCGCGAA 720
Db 3238 GCGTGGCGCCGCAAGATCATCCCTACCAACCCCAAGGCAAGACAGCAGTGTGCGCGAA 3297
OY 721 TACGAGGCGCTCAAGGGCTGGGCCACCGGCACTTGGCCAGCTGCAAGCAGCTACCTC 780
Db 3298 TACGAGGCGCTCAAGGGCTGGGCCACCGGCACTTGGCCAGCTGCAAGCAGCTACCTC 3357
OY 781 AGCCCCCGGCACTGCTCATCTTGGAGCTGTCTTGGCCGAGCTGTCTCCCTGC 840
Db 3358 AGCCCCCGGCACTGCTCATCTTGGAGCTGTCTTGGCCGAGCTGTCTCCCTGC 3417
OY 841 CTGGCCGAGAGGCGCTCTACTCAGATCCGAGGTGAAGGACTACTCTGGCAGATGTTG 900
Db 3418 CTGGCCGAGAGGCGCTCTACTCAGATCCGAGGTGAAGGACTACTCTGGCAGATGTTG 3477
OY 901 AGTGCCACCCAGTACTGCAACAGCAGCATCTTGCACCTGGACCTGAGGTCGAGAAC 960
Db 3478 AGTGCCACCCAGTACTGCAACAGCAGCATCTTGCACCTGGACCTGAGGTCGAGAAC 3537
OY 961 ATGATCATCAGCAATACAACTGCTCAAGTCTGGACCT 1001
Db 3538 ATGATCATCAGCAATACAACTGCTCAAGTCTGGACCT 3578
```

RESULT 3

```
US-10-182-243-14
; Sequence 14, Application US/10182243
; Publication No. US20040048310A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: WHYTE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: MARTINEZ, RICARDO
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE
; FILE REFERENCE: 038602/1366
; CURRENT APPLICATION NUMBER: US/10/182,243
; PRIOR FILING DATE: 2003-07-07
; PRIOR APPLICATION NUMBER: PCT/US01/02337
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
```

```
; LENGTH: 4936
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-243-14

Query Match
Best Local Similarity 100.0%; Score 1001; DB 17; Length 4936;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGCAGTCACTGACACTGGCCCTGCCAGGTGTCCAGCCAGCCAGCTGCCAGGCCACCTGG 60
Db 3379 GGCAGTCACTGACACTGGCCCTGCCAGGTGTCCAGCCAGCCAGCTGCCAGGCCACCTGG 3438
OY 61 AGCAAGAAGAGCCGCCCTCTGAGAGCAGCAGCCCTGTCTCATCTCTGCCACCTCAAG 120
Db 3439 AGCAAGAAGAGCCGCCCTCTGAGAGCAGCAGCCCTGTCTCATCTCTGCCACCTCAAG 3498
OY 121 AACTTCCAGCTTCTGACCATCTGCTGTGTGGTGTGAGGACCTGGGTGTGTACACCTGC 180
Db 3499 AACTTCCAGCTTCTGACCATCTGCTGTGTGGTGTGAGGACCTGGGTGTGTACACCTGC 3558
OY 181 AGCGTGAGCAATGCGCTGGGACAGTGACACCGGCGTCTCTCCGGAAGGACAGAGCG 240
Db 3559 AGCGTGAGCAATGCGCTGGGACAGTGACACCGGCGTCTCTCCGGAAGGACAGAGCG 3618
OY 241 CCCTCATCTTGGCAGTGGCCGATATCGGGAGAGGTGTACGGGATGGGGTGTCTGTGTC 300
Db 3619 CCCTCATCTTGGCAGTGGCCGATATCGGGAGAGGTGTACGGGATGGGGTGTCTGTGTC 3678
OY 301 TGGAAAGCCGCTGGAATCTTACGGCCCTGTGACCTACATTGTGACGTGCAAGCTTGAAGGC 360
Db 3679 TGGAAAGCCGCTGGAATCTTACGGCCCTGTGACCTACATTGTGACGTGCAAGCTTGAAGGC 3738
OY 361 GGCAGCTGGACACACACTGGCCCTCCGACATCTTTTGAATGCTGTCTACTGACCAAGCTC 420
Db 3739 GGCAGCTGGACACACACTGGCCCTCCGACATCTTTTGAATGCTGTCTACTGACCAAGCTC 3798
OY 421 TCCCGGGGTGGCACTCAACCTTCCGCAAGGATGTGTAGCAAGGCAAGGATGGGTGCC 480
Db 3799 TCCCGGGGTGGCACTCAACCTTCCGCAAGGATGTGTAGCAAGGCAAGGATGGGTGCC 3858
OY 481 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGGCCAGCCACTTGGCCCTCTGAGGAG 540
Db 3859 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGGCCAGCCACTTGGCCCTCTGAGGAG 3918
OY 541 GAGAGCCAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTGGCATTCCAGACA 600
Db 3919 GAGAGCCAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTGGCATTCCAGACA 3978
OY 601 CAGATCCAGAGGGCGGCTTACGGTGTGTGGGCAATGCTGGGAGAGGCGCAGCGGCGG 660
Db 3979 CAGATCCAGAGGGCGGCTTACGGTGTGTGGGCAATGCTGGGAGAGGCGCAGCGGCGG 4038
OY 661 GCGCTGGCCGCAAGATCATCCCTACCAACCCCAAGGACAGACAGTGTCTGCGCGAA 720
Db 4039 GCGCTGGCCGCAAGATCATCCCTACCAACCCCAAGGACAGACAGTGTCTGCGCGAA 4098
OY 721 TACGAGGCGCTCAAGGGCTTGGCCACCGCAGCAGCTTGGCCAGCTGACAGCAGCTTACCTC 780
Db 4099 TACGAGGCGCTCAAGGGCTTGGCCACCGCAGCAGCTTGGCCAGCTGACAGCAGCTTACCTC 4158
OY 781 AGCCCCCGGCACTTGGTGTCTCATTTGAGCTGTCTTGGGCCCGGAGTGTCTCCCTGC 840
Db 4159 AGCCCCCGGCACTTGGTGTCTCATTTGAGCTGTCTTGGGCCCGGAGTGTCTCCCTGC 4218
OY 841 CTGGCCGAGAGGGCTCTCTACTCAGATCCGAGGTGAAGGACTACTCTGGCAGATGTTG 900
Db 4219 CTGGCCGAGAGGGCTCTCTACTCAGATCCGAGGTGAAGGACTACTCTGGCAGATGTTG 4278
OY 901 AGTGCCACCCAGTACTGCAACAGCAGCATCTTGCACCTGGAGCTGAGGTCGAGAAC 960
Db 4279 AGTGCCACCCAGTACTGCAACAGCAGCATCTTGCACCTGGAGCTGAGGTCGAGAAC 4338
```


Qy 961 ATGATCATCCGGAATACAACTGCTCAAGTCTGAGACCT 1001
Db 4339 ATGATCATCCGGAATACAACTGCTCAAGTCTGAGACCT 4379

RESULT 4

US-10-415-011-44
; Sequence 44, Application US/10415011
; Publication No. US20040053394A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: CHAWLA, Narinder K.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: XU, Yuming
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: DING, Li
; APPLICANT: TANG, Y. Tom
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LU, Yan
; APPLICANT: YUE, Henry
; APPLICANT: BURFORD, Neil
; APPLICANT: BANDMAN, Olga
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: LAL, Preeti G.
; APPLICANT: REGIPON, Shirley A.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: SWARNAKER, Anita
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: KHAN, Farrah A.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0262 USN
; CURRENT APPLICATION NUMBER: US/10/415,011
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: PCT/US01/47728
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/242,410
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/244,068
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/245,708
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/247,672
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,565
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,730
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,807
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 44
; LENGTH: 5454
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20040053394A1 7638121CB1
US-10-415-011-44

Query Match 100.0%; Score 1001; DB 17; Length 5454;
Best Local Similarity 100.0%; Pred. No. 2.3e-261;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCAGTCACTGACACTGGCCTGCCAGGTCTACGCCAGCCAGCTGCCAGGGCCACTGG 60
Db 3892 GGCAGTCACTGACACTGGCCTGCCAGGTCTGCCAGCCAGCCAGCTGCCAGGGCCACTGG 3951
Qy 61 AGCAAGACGGAGCCCCCTTGGAGAGCAGCAGCGGTCTCTCATCTCTGCCACCTCAAG 120
Db 3952 AGCAAGACGGAGCCCCCTTGGAGAGCAGCAGCGGTCTCTCATCTCTGCCACCTCAAG 4011
Qy 121 AACTTCCAGCTTCTGACCATCTCTGGTGGTGGCTGAGCACTTGGGTGTGTACCTGC 180
Db 4012 AACTTCCAGCTTCTGACCATCTCTGGTGGTGGCTGAGCACTTGGGTGTGTACCTGC 4071
Qy 181 AGCGTGAGCAATCGCTGGGGACAGTCAACCAACCGGGCTCTCCGGAAGCAGAGCGC 240
Db 4072 AGCGTGAGCAATCGCTGGGGACAGTCAACCAACCGGGCTCTCCGGAAGCAGAGCGC 4131
Qy 241 CCCTCATCTTCGCATGCCCGGATATCGGGAGGTGTACCGGATGGGTGCTGCTGTC 300
Db 4132 CCCTCATCTTCGCATGCCCGGATATCGGGAGGTGTACCGGATGGGTGCTGCTGTC 4191
Qy 301 TGGAGCCCGTGGAAATCCTACGGCCCTGTGACCTTACATTTGTGCAAGTGCAGCTAGAAGC 360
Db 4192 TGGAGCCCGTGGAAATCCTACGGCCCTGTGACCTTACATTTGTGCAAGTGCAGCTAGAAGC 4251
Qy 361 GGCAGTGGACCAACATCTGGGCTTCGCAATCTTTTGA CTGCTGCTACCTGACAGCAAGCTC 420
Db 4252 GGCAGTGGACCAACATCTGGGCTTCGCAATCTTTTGA CTGCTGCTACCTGACAGCAAGCTC 4311
Qy 421 TCCCGGGTGGCACCTTACACCTTCCGACAGGCACTGTGTAGCAAGGAGGAATGGGTCCC 480
Db 4312 TCCCGGGTGGCACCTTACACCTTCCGACAGGCACTGTGTAGCAAGGAGGAATGGGTCCC 4371
Qy 481 TACAGCAGCCCTCGAGCAAGTCTCTCGGAGGGCCAGCCACCTTGGCTCTTGAGGAG 540
Db 4372 TACAGCAGCCCTCGAGCAAGTCTCTCGGAGGGCCAGCCACCTTGGCTCTTGAGGAG 4431
Qy 541 GAGAGCAGGGGGCTGAGCCCAACCTTCCGACAGCAAGAGCTTGGCATTTCCAGACA 600
Db 4432 GAGAGCAGGGGGCTGAGCCCAACCTTCCGACAGCAAGAGCTTGGCATTTCCAGACA 4491
Qy 601 CAGATCCAGAGGGGCGCTTTCAGCGTGTGGGCAATGTCTGGGAGAAAGCCAGCGGGCGG 660
Db 4492 CAGATCCAGAGGGGCGCTTTCAGCGTGTGGGCAATGTCTGGGAGAAAGCCAGCGGGCGG 4551
Qy 661 GCGCTGGCCGCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGTCTGCGGAA 720
Db 4552 GCGCTGGCCGCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGTCTGCGGAA 4611
Qy 721 TACAGGCGCTCAAGGGCTTGGCCACCGCACTTGGCCAGCTGCACGAGCTTACCTC 780
Db 4612 TACAGGCGCTCAAGGGCTTGGCCACCGCACTTGGCCAGCTGCACGAGCTTACCTC 4671
Qy 781 AGCCCCCGGCACTTGGTGTCTCATCTTGAGCTGTGCTCTGGGCCCCGAGCTGTCTCCCTGC 840
Db 4672 AGCCCCCGGCACTTGGTGTCTCATCTTGAGCTGTGCTCTGGGCCCCGAGCTGTCTCCCTGC 4731
Qy 841 CTGSCCGAGAGGGCTCTCTACTCAGAAATCCAGGTGAAGGACTTCTGTGGCAGATGTTG 900
Db 4732 CTGSCCGAGAGGGCTCTCTACTCAGAAATCCAGGTGAAGGACTTCTGTGGCAGATGTTG 4791
Qy 901 AGTGCCACCCAGTACCTGCACAAACAGCACATCTTGCACTTGGAGCTTCCGAGAAC 960
Db 4792 AGTGCCACCCAGTACCTGCACAAACAGCACATCTTGCACTTGGAGCTTCCGAGAAC 4851
Qy 961 ATGATCATCCCGGAATACAACTGCTCAAGTCTGAGGACCT 1001
Db 4852 ATGATCATCCCGGAATACAACTGCTCAAGTCTGAGGACCT 4892

RESULT 5
US-10-077-130-3
; Sequence 3, Application US/10077130
; Publication No. US20020168742A1


```
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; FILE REFERENCE: MPI2001-047P1RCPI(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 7893
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-077-130-3

Query Match      100.0%; Score 1001; DB 13; Length 7893;
Best Local Similarity 100.0%; Pred. No. 2.4e-261;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTCACTGACACTGCGCTGCCAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTGG 60
DB 6415 GGCAGTCACTGACACTGCGCTGCCAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTGG 6474

QY 61 AGCAAGACGGAGCCCTCCAGAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTCAAG 120
DB 6475 AGCAAGACGGAGCCCTCCAGAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTCAAG 6534

QY 121 AACTTCCAGTTCAGACATCTGTCAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTGC 180
DB 6535 AACTTCCAGTTCAGACATCTGTCAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTGC 6594

QY 181 AGCGTGAGCAATCGCTGGGACAGTACACCAAGCGGCTCTCCGGAAGGACAGAGCG 240
DB 6595 AGCGTGAGCAATCGCTGGGACAGTACACCAAGCGGCTCTCCGGAAGGACAGAGCG 6654

QY 241 CCTCATCTTCGCATGTCGGAGAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTCAAG 300
DB 6655 CCTCATCTTCGCATGTCGGAGAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTCAAG 6714

QY 301 TGAAGCCGCTGGAATCTTACCGCCCTGTGACCTACATTTGACTGTCTACCAAGAGCTC 360
DB 6715 TGAAGCCGCTGGAATCTTACCGCCCTGTGACCTACATTTGACTGTCTACCAAGAGCTC 6774

QY 361 GGCAGTGGACACACTGCGCTCCGACATCTTTGACTGTCTACCTACCAAGAGCTC 420
DB 6775 GGCAGTGGACACACTGCGCTCCGACATCTTTGACTGTCTACCTACCAAGAGCTC 6834

QY 421 TCCGGGTTGGACCTACACCTTCGCGACGGCATGTGTGAGCAAGGAGGAAATGGGTCCC 480
DB 6835 TCCGGGTTGGACCTACACCTTCGCGACGGCATGTGTGAGCAAGGAGGAAATGGGTCCC 6894

QY 481 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGGCCAGCACTGGCCCTCTGAGGAG 540
DB 6895 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGGCCAGCACTGGCCCTCTGAGGAG 6954

QY 541 GAGAGCCAGGGCGGTTCAGCCCAAGCCCTCGCCAGCACAAGACCTTCGATTTCCAGACA 600
DB 6955 GAGAGCCAGGGCGGTTCAGCCCAAGCCCTCGCCAGCACAAGACCTTCGATTTCCAGACA 7014

QY 601 CAGATCCAGAGGGCGGTTCAGCGTGTGCGGCAATGCTGGGAGAAAGCCAGCGGGCGG 660
DB 7015 CAGATCCAGAGGGCGGTTCAGCGTGTGCGGCAATGCTGGGAGAAAGCCAGCGGGCGG 7074

QY 661 GCGTGGCGCCCAAGATCATCCCTACCAAGCCAGGACAGCAGTGTGTCGCGGAA 720
DB 7075 GCGTGGCGCCCAAGATCATCCCTACCAAGCCAGGACAGCAGTGTGTCGCGGAA 7134

QY 721 TACAGAGGCGCTCAAGGGCTGGCCCAAGCCAGCTGGCCAGCTGGCCAGCTCACTC 780
DB 7135 TACAGAGGCGCTCAAGGGCTGGCCCAAGCCAGCTGGCCAGCTGGCCAGCTCACTC 7194

; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; FILE REFERENCE: MPI2001-047P1RCPI(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 8106
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: 5' UTR
; LOCATION: (1) ... (71)
; NAME/KEY: CDS
; LOCATION: (72) ... (7964)
; NAME/KEY: 3' UTR
; LOCATION: (7965) ... (8106)
US-10-077-130-1

Query Match      100.0%; Score 1001; DB 13; Length 8106;
Best Local Similarity 100.0%; Pred. No. 2.4e-261;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCAGTCACTGACACTGCGCTGCCAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTGG 60
DB 6486 GGCAGTCACTGACACTGCGCTGCCAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTGG 6545

QY 61 AGCAAGACGGAGCCCTCCAGAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTCAAG 120
DB 6546 AGCAAGACGGAGCCCTCCAGAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTCAAG 6605

QY 121 AACTTCCAGTTCAGACATCTGTCAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTGC 180
DB 6606 AACTTCCAGTTCAGACATCTGTCAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTGC 6665

QY 181 AGCGTGAGCAATCGCTGGGACAGTACACCAAGCGGCTCTCCGGAAGGACAGAGCG 240
DB 6666 AGCGTGAGCAATCGCTGGGACAGTACACCAAGCGGCTCTCCGGAAGGACAGAGCG 6725

QY 241 CCTCATCTTCGCATGTCGGAGAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTCAAG 300
DB 6726 CCTCATCTTCGCATGTCGGAGAGTGTGTCAGCCAGCCAGCTGCCAGGCCACCTCAAG 6785

QY 301 TGAAGCCGCTGGAATCTTACCGCCCTGTGACCTACATTTGACTGTCTACCAAGAGCTC 360
DB 7135 TGAAGCCGCTGGAATCTTACCGCCCTGTGACCTACATTTGACTGTCTACCAAGAGCTC 7194
```


6786 TGGAAGCCCGTGGAAATCCTTAGCGCCCTGTGACCTACATTGTGCACTGCGAGCTAGAGGC 6845
Db
361 GGCAGCTGGACCAACACTGGCCCTCGACATCTTTTGACTGCTGCTACTGACAGCAAGCTC 420
Qy
6846 GGCAGCTGGACCAACACTGGCCCTCGACATCTTTTGACTGCTGCTACTGACAGCAAGCTC 6905
Db
421 TCCCGGGGTGGCACTTACACCTTCGGCAACGCGATGTGTGACAGGAGGAGGAGTGGTCCC 480
Qy
6906 TCCCGGGGTGGCACTTACACCTTCGGCAACGCGATGTGTGACAGGAGGAGGAGTGGTCCC 6965
Db
481 TACAGCAGCCCTCGAGCAAGTCTCTCTGGAGAGGCGCCAGCCACCTGGCCCTCTGAGGAG 540
Qy
6966 TACAGCAGCCCTCGAGCAAGTCTCTCTGGAGAGGCGCCAGCCACCTGGCCCTCTGAGGAG 7025
Db
541 GAGAGCCAGGGGCGGTGACGCCCAACCCCTGCGCCAGCAACAAAGACTTTCGCAATTCAGACA 600
Qy
7026 GAGAGCCAGGGGCGGTGACGCCCAACCCCTGCGCCAGCAACAAAGACTTTCGCAATTCAGACA 7085
Db
601 CAGATCCAGAGGGGCGGTTCAGCGGTGTGGGCAATGCTGGAGAGGCGCCAGCGGCGG 660
Qy
7086 CAGATCCAGAGGGGCGGTTCAGCGGTGTGGGCAATGCTGGAGAGGCGCCAGCGGCGG 7145
Db
661 GCGCTGGCGCCCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGCAGTGTCTGCGGAA 720
Qy
7146 GCGCTGGCGCCCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGCAGTGTCTGCGGAA 7205
Db
721 TACAGAGCCCTCAAGGGCGTGGCCACCGGCACTGGCCAGCTGCGAGCCCTACTCTC 780
Qy
7206 TACAGAGCCCTCAAGGGCGTGGCCACCGGCACTGGCCAGCTGCGAGCCCTACTCTC 7265
Db
781 AGCCCGCGCACCTGGTGTCTATCTTGGAGCTGTGCTCTGGGCCCGAGCTGTCTCCCTGC 840
Qy
7266 AGCCCGCGCACCTGGTGTCTATCTTGGAGCTGTGCTCTGGGCCCGAGCTGTCTCCCTGC 7325
Db
841 CTGGCGGAGAGGGCCCTCTACTCAGAAATCCGAGGTGAAGGACTACTGTGGCAGATGTTG 900
Qy
7326 CTGGCGGAGAGGGCCCTCTACTCAGAAATCCGAGGTGAAGGACTACTGTGGCAGATGTTG 7385
Db
901 AGTGCCACCCAGTACTGACAGCAGCAGCATCTGCACTGCACTGAGGTCCGAGAAC 960
Qy
7386 AGTGCCACCCAGTACTGACAGCAGCAGCATCTGCACTGCACTGAGGTCCGAGAAC 7445
Db
961 ATGATCATCACCAGATACAACTGTCTCAAGGTCTGGACCT 1001
Qy
7446 ATGATCATCACCAGATACAACTGTCTCAAGGTCTGGACCT 7486
Db

RESULT 7

US-10-077-130-6
; Sequence 6, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047P1RCP1(W)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR FILING DATE: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 23907
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-077-130-6

Query Match 100.0%; Score 1001; DB 13; Length 23907;
Best Local Similarity 100.0%; Pred. No. 2,7e-261;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 8

US-10-077-130-4
; Sequence 4, Application US/10077130
; Publication No. US20020168742A1

1 GGCCAGTCACTGACACTGGCCCTGCCAGTGTGAGCCACGACGCTGCCAGGCGACCTGG 60
Db
22429 GGCCAGTCACTGACACTGGCCCTGCCAGTGTGAGCCACGACGCTGCCAGGCGACCTGG 22488
Qy
61 AGCAAGAAGAGCGAGCCCTCGAGAGCAGAGCGGTGCTCATCTCTGCCACCTCAAG 120
Db
22489 AGCAAGAAGAGCGAGCCCTCGAGAGCAGAGCGGTGCTCATCTCTGCCACCTCAAG 22548
Qy
121 AACTTCCAGTCTTGTACCATCTCTGGTGTGTGGCTGAGGACCTGGGTGTGTACCTGC 180
Db
22549 AACTTCCAGTCTTGTACCATCTCTGGTGTGTGGCTGAGGACCTGGGTGTGTACCTGC 22608
Qy
181 AGCGTAGCAATGCGTGGGGACAGTGACCAACACGCGGGGTCTCTCGGAAGGAGAGCGC 240
Db
22609 AGCGTAGCAATGCGTGGGGACAGTGACCAACACGCGGGGTCTCTCGGAAGGAGAGCGC 22668
Qy
241 CCTCATCTTCGCGATGCGCGGATATCGGGAGGTGTAGCGGATGGGGTGTCTGTGGTC 300
Db
22669 CCTCATCTTCGCGATGCGCGGATATCGGGAGGTGTAGCGGATGGGGTGTCTGTGGTC 22728
Qy
301 TGAAGCCCGTGGAAATCTACGGCCCTGTGACCTACATTGTGCACTGAGCCTAGAGGC 360
Db
22729 TGAAGCCCGTGGAAATCTACGGCCCTGTGACCTACATTGTGCACTGAGCCTAGAGGC 22788
Qy
361 GGCACTGGACCAACACTGGGCTCGGACATCTTTGACTGTGCTACTGACAGCAAGCTC 420
Db
22789 GGCACTGGACCAACACTGGGCTCGGACATCTTTGACTGTGCTACTGACAGCAAGCTC 22848
Qy
421 TCCGGGGTGGCACTTACCTTCCGCAACGCGATGTGTGAGCAAGGAGGAGTGGTCCC 480
Db
22849 TCCGGGGTGGCACTTACCTTCCGCAACGCGATGTGTGAGCAAGGAGGAGTGGTCCC 22908
Qy
481 TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGGCCAGCCACCTGGGCTCTGAGGAG 540
Db
22909 TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGGCCAGCCACCTGGGCTCTGAGGAG 22968
Qy
541 GAGAGCAGAGGGGCGGTGAGCCCAACCCCTGCCAGCACAAGACTTTCGATTCAGACA 600
Db
22969 GAGAGCAGAGGGGCGGTGAGCCCAACCCCTGCCAGCACAAGACTTTCGATTCAGACA 23028
Qy
601 CAGATCCAGAGGGGCGGTGAGCCCAACCCCTGCCAGCACAAGACTTTCGATTCAGACA 660
Db
23029 CAGATCCAGAGGGGCGGTGAGCCCAACCCCTGCCAGCACAAGACTTTCGATTCAGACA 23088
Qy
661 GCGTGGCGCCAAAGATCATCCCTTACCAACCCCAAGGACAGAGCAGAGTGTGCGGAA 720
Db
23089 GCGTGGCGCCAAAGATCATCCCTTACCAACCCCAAGGACAGAGCAGAGTGTGCGGAA 23148
Qy
721 TACGAGGCGCTCAAGGGCTGCGGCACCCGACCTGCGCCAGCTGCGACAGCCTTACCTC 780
Db
23149 TACGAGGCGCTCAAGGGCTGCGGCACCCGACCTGCGCCAGCTGCGACAGCCTTACCTC 23208
Qy
781 AGCCCGCGCACCTGGTGTCTCATCTTGGAGTGTGCTCTGGGCCCGAGCTGTCTCCCTGC 840
Db
23209 AGCCCGCGCACCTGGTGTCTCATCTTGGAGTGTGCTCTGGGCCCGAGCTGTCTCCCTGC 23268
Qy
841 CTGGCCGAGAGGGCCCTCTACTCAGAAATCCGAGGTGAAGGACTACTGTGGCAGATGTTG 900
Db
23269 CTGGCCGAGAGGGCCCTCTACTCAGAAATCCGAGGTGAAGGACTACTGTGGCAGATGTTG 23328
Qy
901 AGTGCCACCCAGTACTGCAACCAACAGCAGCATCTCTGCACTTGGACCTGAGGTCCGAGAAC 960
Db
23329 AGTGCCACCCAGTACTGCAACCAACAGCAGCATCTCTGCACTTGGACCTGAGGTCCGAGAAC 23388
Qy
961 ATGATCATCACCAGATACAACTGTCTCAAGGTCTGGACCT 1001
Db
23389 ATGATCATCACCAGATACAACTGTCTCAAGGTCTGGACCT 23429


```

; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047P1RCPI(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 24120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(71)
; NAME/KEY: CDS
; LOCATION: (72)...(23978)
; NAME/KEY: 3'UTR
; LOCATION: (23979)...(24120)
US-10-077-130-4

```

Query Match	100.00%;	Score 1001;	DB 13;	Length 24120;
Best Local Similarity	100.0%;	Pred. No. 2.7e-261;		
Matches 1001;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GGCCAGTCAGTGACACTGCGCTGCCAGAGGTGTCAGCCACGACAGCTGCCACGGCCACCTGG	60	
Db	23500	GGCCAGTCAGTGACACTGCGCTGCCAGAGGTGTCAGCCACGACAGCTGCCACGGCCACCTGG	22559	
Qy	61	AGCAAAAGACGGAGCCCCCTCGAGAGACGACGCGGTCTCATCTCTGCCACCCCTCAAG	120	
Db	22560	AGCAAAAGACGGAGCCCCCTCGAGAGACGACGCGGTCTCATCTCTGCCACCCCTCAAG	22619	
Qy	121	AAC TTCAGCTTCTGACCATCTGTGTGGTGCTGGCTGAGGACCTGGGTGTGTACACCTGC	180	
Db	22620	AAC TTCAGCTTCTGACCATCTGTGTGGTGCTGGCTGAGGACCTGGGTGTGTACACCTGC	22679	
Qy	181	AGCGTGAGCAATGCGCTGGGGACAGTGACCAACA GGGCGTCTCTCGGAAGGACGAGCGC	240	
Db	22680	AGCGTGAGCAATGCGCTGGGGACAGTGACCAACA GGGCGTCTCTCGGAAGGACGAGCGC	22739	
Qy	241	CCCTCATCTTCGCGCATGCCCGGATATCGGGGAGGTGTACGGGATGGGGTGTCTGTGTC	300	
Db	22740	CCCTCATCTTCGCGCATGCCCGGATATCGGGGAGGTGTACGGGATGGGGTGTCTGTGTC	22799	
Qy	301	TGGAAGCCCGTGGAATCTTACGGCCCTGTGACCTACTATTGTGCAGCTGACGCTTAGAAGGC	360	
Db	22800	TGGAAGCCCGTGGAATCTTACGGCCCTGTGACCTACTATTGTGCAGCTGACGCTTAGAAGGC	22859	
Qy	361	GGCAGCTGGAACCAACACTGCGCTCCGACATCTTTTGA CTGCTGCTACTCACTCAGCAAGCTC	420	
Db	22860	GGCAGCTGGAACCAACACTGCGCTCCGACATCTTTTGA CTGCTGCTACTCACTCAGCAAGCTC	22919	
Qy	421	TCCCGGGGTGGCACTTACACCTTCGCGACGGCATGTGTACGAAAGGACAGGAATGGGTCCC	480	
Db	22920	TCCCGGGGTGGCACTTACACCTTCGCGACGGCATGTGTACGAAAGGACAGGAATGGGTCCC	22979	
Qy	481	TACAGACGCCCTTCGGAGCAAGTCTCTCTGGAGGGGCCAGGCACCTTGGCTCTGAGGAG	540	
Db	22980	TACAGACGCCCTTCGGAGCAAGTCTCTCTGGAGGGGCCAGGCACCTTGGCTCTGAGGAG	23039	
Qy	541	GAGAGCCAGGGGCGGTTCAGCCCAACCCCTGCCAGCAAAAGACCTTCGCAATCCAGACA	600	
Db	23040	GAGAGCCAGGGGCGGTTCAGCCCAACCCCTGCCAGCAAAAGACCTTCGCAATCCAGACA	23099	
Qy	601	CAGATCCAGGGGCGGCTTACGCTGTGTGGGCAATGCTGGGAGAAAGCCAGCGGGCGG	660	
Db	23100	CAGATCCAGGGGCGGCTTACGCTGTGTGGGCAATGCTGGGAGAAAGCCAGCGGGCGG	23159	

Qy	661	GGCTGGCGCGCAAGATCATCCCTACCAACCCCAAGGACAAGACAGCAGTGTCTGGCGAA	720
Db	23160	GGCTGGCGCGCAAGATCATCCCTACCAACCCCAAGGACAAGACAGCAGTGTCTGGCGAA	23219
Qy	721	TACGAGGCCCTCAAGGGCTGGCCACCCGACACCTTGGCCAGCTGCACGACGCTTACCTC	780
Db	23220	TACGAGGCCCTCAAGGGCTGGCCACCCGACACCTTGGCCAGCTGCACGACGCTTACCTC	23279
Qy	781	AGCCCCCGGCACCTGGTGTCTCATCTTGGAGCTGTGCTCTGGGCCCGAGCTGTCCCTTCG	840
Db	23280	AGCCCCCGGCACCTGGTGTCTCATCTTGGAGCTGTGCTCTGGGCCCGAGCTGTCCCTTCG	23339
Qy	841	CTGGCCGAGAGGGCTCTCTACTCAGAAATCCGAGGTGAAGGACTACCTGTGGCAGATGTTG	900
Db	23340	CTGGCCGAGAGGGCTCTCTACTCAGAAATCCGAGGTGAAGGACTACCTGTGGCAGATGTTG	23399
Qy	901	AGTGCCACCCAGTACTCTGCACAACAGACATCTCTGACCTTGGACCTCCGAGAAC	960
Db	23400	AGTGCCACCCAGTACTCTGCACAACAGACATCTCTGACCTTGGACCTCCGAGAAC	23459
Qy	961	ATGATCATCACCGAATACAACTGCTCAAGGTCGTGGACCT	1001
Db	23460	ATGATCATCACCGAATACAACTGCTCAAGGTCGTGGACCT	23500

RESULT 9

US-10-307-019-3

; Sequence 3, Application US/10307019

; Publication No. US20030108533A1

; GENERAL INFORMATION:

; APPLICANT: Zeng, Wenlin

; APPLICANT: Stanton, Lawrence

; APPLICANT: SCIOS, INC.

; TITLE OF INVENTION: MEDIATORS OF SIGNAL TRANSDUCTION

; FILE REFERENCE: SCIOS.021DV1

; CURRENT APPLICATION NUMBER: US/10/307,019

; CURRENT FILING DATE: 2002-11-26

; PRIOR APPLICATION NUMBER: 09/548,473

; PRIOR FILING DATE: 2000-04-13

; PRIOR APPLICATION NUMBER: 60/129,552

; PRIOR FILING DATE: 1999-04-16

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 3

; LENGTH: 5007

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (97) ... (4926)

US-10-307-019-3

Query Match	99.8%;	Score 999.4;	DB 15;	Length 5007;
Best Local Similarity	99.9%;	Pred. No. 6.3e-261;		
Matches 1000;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	GGCGAGTCACTGACACTGGCCCTGCGAGGTGTACGCCACGAGTGTGCCAGGCACCTGG	60	
Db	3451	GGCCAGTCACTGACACTGGCCCTGCCAGGTGTACGCCACGAGTGTGCCAGGCACCTGG	3510	
Qy	61	AGCAAGACGGAGCCGCCCTTGAGAGCAGCAGCCGTGTCTCATCTCTGCACACCTCAAG	120	
Db	3511	AGCAAGACGGAGCCGCCCTTGAGAGCAGCAGCCGTGTCTCATCTCTGCACACCTCAAG	3570	
Qy	121	AAC TTCAGCTTCTGACCATCTCGTGTGTGGCTGAGGACCTGGGTGTGTACACCTGC	180	
Db	3571	AAC TTCAGCTTCTGACCATCTCGTGTGTGGCTGAGGACCTGGGTGTGTACACCTGC	3630	
Qy	181	AGCGTAGCAATGCGCTGGGACAGTACACCAACCGGCGTCTCTCCGGAAGCAGAGCGC	240	
Db	3631	AGCGTAGCAATGCGCTGGGACAGTACACCAACCGGCGTCTCTCCGGAAGCAGAGCGC	3690	
Qy	241	CCCTCATCTTCGCCATGCCCAGATATCGGGAGAGGTGTACGGCGATGGGGTGTCTGTGGTC	300	

Qy 301 TGGAGGCCGCTGGGAATCCTACGGCCCTGTGACCTACATTTGTGCAGTGCAGCCTAGAAGGC 360
Db |||||
Qy 3928 TGGAGGCCGCTGGGAATCCTACGGCCCTGTGACCTACATTTGTGCAGTGCAGCCTAGAAGGC 3987
Db |||||
Qy 361 GGCAGCTGGACCACTAGGCTCTCGACATCTTTGACTGTCTGCTACTCTGACCAAGCTC 420
Db |||||
Qy 3988 GGCAGCTGGACCACTAGGCTCTCGACATCTTTGACTGTCTGCTACTCTGACCAAGCTC 4047
Db |||||
Qy 421 TCCCGGGGTGGCACTACACCTTCGGACGCGATGTGTGACAAAGGAGGAATGGTCCC 480
Db |||||
Qy 4048 TCCCGGGGTGGCACTACACCTTCGGACGCGATGTGTGACAAAGGAGGAATGGTCCC 4107
Db |||||
Qy 481 TACAGACGCCCTCGGAGCAAGTCTCTCGGAGGGGCCAGCCACTCTGAGGAG 540
Db |||||
Qy 4108 TACAGACGCCCTCGGAGCAAGTCTCTCGGAGGGGCCAGCCACTCTGAGGAG 4167
Db |||||
Qy 541 GAGAGCCAGGGGGGTGAGCCCAACCCCTGCTGCCAGCAAAAGACTTTCGATTCGAGACA 600
Db |||||
Qy 4168 GAGAGCCAGGGGGGTGAGCCCAACCCCTGCTGCCAGCAAAAGACTTTCGATTCGAGACA 4227
Db |||||
Qy 601 CAGATCCAGAGGGGGCCGCTTCAGCGTGTGTGGGCAATGTCTGGGAGAGGCCAGCGGGCGG 660
Db |||||
Qy 4228 CAGATCCAGAGGGGGCCGCTTCAGCGTGTGTGGGCAATGTCTGGGAGAGGCCAGCGGGCGG 4287
Db |||||
Qy 661 GCGTGGCCGCGCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGCAGTGTCTGCGGAA 720
Db |||||
Qy 4288 GCGTGGCCGCGCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGCAGTGTCTGCGGAA 4347
Db |||||
Qy 721 TACAGAGCCCTCAAGGGCTCGGCGACCCGACCTGGGCCAGCTGCGAGCCTACTCTC 780
Db |||||
Qy 4348 TACAGAGCCCTCAAGGGCTCGGCGACCCGACCTGGGCCAGCTGCGACAGCCTACTCTC 4407
Db |||||
Qy 781 AGCCCCGGCACCTGGTGTCTCATCTTGAGAGTGTGTCTCTGGGCCCGAGCTCTCTCCCTGC 840
Db |||||
Qy 4408 AGCCCCGGCACCTGGTGTCTCATCTTGAGAGTGTGTCTCTGGGCCCGAGCTCTCTCCCTGC 4467
Db |||||
Qy 841 CTGGCCGAGAGGGCCCTCTACTCAGAATCCGAGGTGAAGACTACTCTGTGGCAGATGTTG 900
Db |||||
Qy 4468 CTGGCCGAGAGGGCCCTCTACTCAGAATCTGAGGTGAAGACTACTCTGTGGCAGATGTTG 4527
Db |||||
Qy 901 AGTGCCACCCAGTACTGTCACAAACAGACATCTGACCTGGAGCTGAGTCCGAGAAC 960
Db |||||
Qy 4528 AGTGCCACCCAGTACTGTCACAAACAGACATCTGACCTGGAGCTGAGTCCGAGAAC 4587
Db |||||
Qy 961 ATGATCATCCCGAATACAACTGCTCAAGGTCTGGGACCT 1001
Db |||||
Qy 4588 ATGATCATCCCGAATACAACTGCTCAAGGTCTGGGACCT 4628
Db |||||

RESULT 13

US-10-697-263-1
; Sequence 1, Application US/10697263
; Publication No. US20040063142A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,263
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274, 978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858, 664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711, 134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-697-263-1
Query Match 99.8%; Score 999.4; DB 17; Length 5207;
Best Local Similarity 99.9%; Pred. No. 6.3e-261;
Matches 1000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GGCAGTCACTGACACTGGGCTGCGAGGTGTGAGCCAGCAGCTGTGCCAGGCACTCTGG 60
Db |||||
Qy 3628 GGCAGTCACTGACACTGGGCTGCGAGGTGTGAGCCAGCAGCTGTGCCAGGCACTCTGG 3687
Db |||||
Qy 61 AGCAAAACGAGGCCCTTGGAGAGCAGCAGCGTGTCTCTCATCTCTGCCACCTCAAG 120
Db |||||
Qy 3688 AGCAAAACGAGGCCCTTGGAGAGCAGCAGCGTGTCTCTCATCTCTGCCACCTCAAG 3747
Db |||||
Qy 121 AACTTCCAGCTTCTGACATCTCTGGTGTGTGCTGAGGACCTGGGTGTGTACACCTGC 180
Db |||||
Qy 3748 AACTTCCAGCTTCTGACATCTCTGGTGTGTGCTGAGGACCTGGGTGTGTACACCTGC 3807
Db |||||
Qy 181 AGCGTAGCAATGCGCTGGGGACAGTGAACAACGCGGCTCTCCGGAAGGACAGGCGC 240
Db |||||
Qy 3808 AGCGTAGCAATGCGCTGGGGACAGTGAACAACGCGGCTCTCCGGAAGGACAGGCGC 3867
Db |||||
Qy 241 CCCTCATCTTTCGCATGCGCCGATATCGGGGAGGTGTACCGGATGGGTCTCTGCTGC 300
Db |||||
Qy 3868 CCCTCATCTTTCGCATGCGCCGATATCGGGGAGGTGTACCGGATGGGTCTCTGCTGC 3927
Db |||||
Qy 301 TGGAGGCCGCTGGGAATCCTACGGCCCTGTGACCTACATTTGTGCAGTGCAGCCTAGAAGGC 360
Db |||||
Qy 3928 TGGAGGCCGCTGGGAATCCTACGGCCCTGTGACCTACATTTGTGCAGTGCAGCCTAGAAGGC 3987
Db |||||
Qy 361 GGCAGCTGGACCACTAGGCTCTCGACATCTTTGACTGTCTGCTACTCTGACCAAGCTC 420
Db |||||
Qy 3988 GGCAGCTGGACCACTAGGCTCTCGACATCTTTGACTGTCTGCTACTCTGACCAAGCTC 4047
Db |||||
Qy 421 TCCCGGGGTGGCACTACACCTTCGGACGCGATGTGTGACAAAGGAGGAATGGTCCC 480
Db |||||
Qy 4048 TCCCGGGGTGGCACTACACCTTCGGACGCGATGTGTGACAAAGGAGGAATGGTCCC 4107
Db |||||
Qy 481 TACAGACGCCCTCGGAGCAAGTCTCTCTGGAGGGGCCAGCCACTCTGAGGAG 540
Db |||||
Qy 4108 TACAGACGCCCTCGGAGCAAGTCTCTCTGGAGGGGCCAGCCACTCTGAGGAG 4167
Db |||||
Qy 541 GAGAGCCAGGGGGCTGAGCCCAACCCCTGCTGCCAGCAAAAGACTTTCGATTCGAGACA 600
Db |||||
Qy 4168 GAGAGCCAGGGGGCTGAGCCCAACCCCTGCTGCCAGCAAAAGACTTTCGATTCGAGACA 4227
Db |||||
Qy 601 CAGATCCAGAGGGGGCCGCTTCAGCGTGTGTGGGCAATGTCTGGGAGAGGCCAGCGGGCGG 660
Db |||||
Qy 4228 CAGATCCAGAGGGGGCCGCTTCAGCGTGTGTGGGCAATGTCTGGGAGAGGCCAGCGGGCGG 4287
Db |||||
Qy 661 GCGTGGCCGCGCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGCAGTGTCTGCGGAA 720
Db |||||
Qy 4288 GCGTGGCCGCGCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGCAGTGTCTGCGGAA 4347
Db |||||
Qy 721 TACAGAGCCCTCAAGGGCTCGGCGACCCGACCTGGGCCAGCTGCGAGCCTACTCTC 780
Db |||||
Qy 4348 TACAGAGCCCTCAAGGGCTCGGCGACCCGACCTGGGCCAGCTGCGACAGCCTACTCTC 4407
Db |||||
Qy 781 AGCCCCGGCACCTGGTGTCTCATCTTGAGAGTGTGTCTCTGGGCCCGAGCTCTCTCCCTGC 840
Db |||||
Qy 4408 AGCCCCGGCACCTGGTGTCTCATCTTGAGAGTGTGTCTCTGGGCCCGAGCTCTCTCCCTGC 4467
Db |||||
Qy 841 CTGGCCGAGAGGGCCCTCTACTCAGAATCCGAGGTGAAGACTACTCTGTGGCAGATGTTG 900
Db |||||
Qy 4468 CTGGCCGAGAGGGCCCTCTACTCAGAATCTGAGGTGAAGACTACTCTGTGGCAGATGTTG 4527
Db |||||
Qy 901 AGTGCCACCCAGTACTGTCACAAACAGACATCTGACCTGGAGCTGAGTCCGAGAAC 960
Db |||||
Qy 4528 AGTGCCACCCAGTACTGTCACAAACAGACATCTGACCTGGAGCTGAGTCCGAGAAC 4587
Db |||||
Qy 961 ATGATCATCCCGAATACAACTGCTCAAGGTCTGGGACCT 1001
Db |||||
Qy 4588 ATGATCATCCCGAATACAACTGCTCAAGGTCTGGGACCT 4628
Db |||||


```
RESULT 14
US-10-697-263-3
; Sequence 3, Application US/10697263
; Publication No. US20040063142A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL00927-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,263
; PRIOR FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-697-263-3

Query Match          99.8%; Score 999.4; DB 17; Length 5207;
Best Local Similarity 99.9%; Pred. No. 6.3e-261;
Matches 1000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCAGTCAGTGACACTGGCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGG 60
DB 3628 GGCAGTCAGTGACACTGGCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGG 3687
QY 61 AGCAAGACGGAGCCCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGG 120
DB 3688 AGCAAGACGGAGCCCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGG 3747
QY 121 AACTTCCAGCTTCTGACCATCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGG 180
DB 3748 AACTTCCAGCTTCTGACCATCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGG 3807
QY 181 AGCTGAGCAATCGCTGGGACAGTACACACAGCGGGTCTCCGGAAGCAGAGCGC 240
DB 3808 AGCTGAGCAATCGCTGGGACAGTACACACAGCGGGTCTCCGGAAGCAGAGCGC 3867
QY 241 CCCTCATCTTCGCCATGCCGATATCGGGAGGTGTACGGGATGGGGTGTCTGGTTC 300
DB 3868 CCCTCATCTTCGCCATGCCGATATCGGGAGGTGTACGGGATGGGGTGTCTGGTTC 3927
QY 301 TGGAGCCCGTGGAAATCTAGCGCCCTGTGACCTATGTCAGTGTGAGCCAGCTAGAGGC 360
DB 3928 TGGAGCCCGTGGAAATCTAGCGCCCTGTGACCTATGTCAGTGTGAGCCAGCTAGAGGC 3987
QY 361 GGCAGTCGACACACTGGCTGCGACATCTTTGACTGCTGCTACCTGACCCAGCAAGTTC 420
DB 3988 GGCAGTCGACACACTGGCTGCGACATCTTTGACTGCTGCTACCTGACCCAGCAAGTTC 4047
QY 421 TCCCGGGTGGACCTACACCTTCCGACGCGATGTGTGAGCCAGCCAGGAGGATGGGTCC 480
DB 4048 TCCCGGGTGGACCTACACCTTCCGACGCGATGTGTGAGCCAGCCAGGAGGATGGGTCC 4107
QY 481 TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCAGCTGCGCTCTGAGGAG 540
DB 4108 TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCAGCTGCGCTCTGAGGAG 4167
QY 541 GAGAGCCAGGGCGGTGACGCCCAACCCCTGCGCCAGCACAAGACCTTTCGCAATCCAGACA 600
DB 4168 GAGAGCCAGGGCGGTGACGCCCAACCCCTGCGCCAGCACAAGACCTTTCGCAATCCAGACA 4227
QY 601 CAGATCCAGAGGGGCGGCTTCAGCGTGTGGGCAATCTCTGGGAGAGGCGGCGGCGG 660
```

```
DB 4228 CAGATCCAGAGGGGCGGCTTCAGCGTGTGGGCAATCTCTGGGAGAGGCGGCGGCGG 4287
QY 661 GCGCTGGCGGCAAGATCATCCCTTACCAACCCCAAGAGCAAGAGCAGTGTCTGGCGGAA 720
DB 4288 GCGCTGGCGGCAAGATCATCCCTTACCAACCCCAAGAGCAAGAGCAGTGTCTGGCGGAA 4347
QY 721 TACGAGGCGCTCAAGGGCTGCGGACCCGACACCTGCGCCAGCTGCGCCAGCTACCTC 780
DB 4348 TACGAGGCGCTCAAGGGCTGCGGACCCGACACCTGCGCCAGCTGCGCCAGCTACCTC 4407
QY 781 AGCCCCCGGACCTTGTGCTCATCTTGGAGTGTGCTCTGGGCGGAGTGTCTCCCTGC 840
DB 4408 AGCCCCCGGACCTTGTGCTCATCTTGGAGTGTGCTCTGGGCGGAGTGTCTCCCTGC 4467
QY 841 CTGGCCGAGAGGGCTCTCTACTCAGATCCGAGTGAAGGACTACCTGTGCGAGATGTTG 900
DB 4468 CTGGCCGAGAGGGCTCTCTACTCAGATCCGAGTGAAGGACTACCTGTGCGAGATGTTG 4527
QY 901 AGTGCCACCCAGTACCTGACAAACAGCAGCATCTGACCTGGACCTGAGGTCGAGAAC 960
DB 4528 AGTGCCACCCAGTACCTGACAAACAGCAGCATCTGACCTGGACCTGAGGTCGAGAAC 4587
QY 961 ATGATCATCACCGAATACAACTGCTCAAGGTCTGTGGACCT 1001
DB 4588 ATGATCATCACCGAATACAACTGCTCAAGGTCTGTGGACCT 4628
```

```
RESULT 15
US-10-921-168-1
; Sequence 1, Application US/10921168
; Publication No. US20050003446A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL00927-CIP-DIV3
; CURRENT APPLICATION NUMBER: US/10/921,168
; CURRENT FILING DATE: 2004-08-19
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-921-168-1
```

```
Query Match          99.8%; Score 999.4; DB 18; Length 5207;
Best Local Similarity 99.9%; Pred. No. 6.3e-261;
Matches 1000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCAGTCAGTGACACTGGCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGG 60
DB 3628 GGCAGTCAGTGACACTGGCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTGG 3687
QY 61 AGCAAGACGGAGCCCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTCAAG 120
DB 3688 AGCAAGACGGAGCCCTGCGAGTGTGAGCCAGCCAGCTGCGCCAGCCACCTCAAG 3747
QY 121 AACTTCCAGCTTCTGACCATCTGCGAGTGTGAGCCAGCCAGCTGCGCTGTACACCTGC 180
DB 3748 AACTTCCAGCTTCTGACCATCTGCGAGTGTGAGCCAGCCAGCTGCGCTGTACACCTGC 3807
QY 181 AGCTGAGCAATCGCTGGGACAGTGTACACACAGCGGGTCTCCGGAAGCAGAGCGC 240
DB 3808 AGCTGAGCAATCGCTGGGACAGTGTACACACAGCGGGTCTCCGGAAGCAGAGCGC 3867
```



```
Qy 241 CCCTCATCTTCGCATGCCCGGATATCGGGAGGTGTACGGGATGGGTGCTGCTGTC 300
Db |||
Qy 3868 CCCTCATCTTCGCATGCCCGGATATCGGGAGGTGTACGGGATGGGTGCTGCTGTC 3927
Db |||
Qy 301 TGAAGCCCGTGGAAATCCTACGGCCCTGTGACCTTACATTGTGCAAGTGCAGCCTAGAAGGC 360
Db |||
Qy 3928 TGAAGCCCGTGGAAATCCTACGGCCCTGTGACCTTACATTGTGCAAGTGCAGCCTAGAAGGC 3987
Db |||
Qy 361 GGCAGCTGGACCACTGGCCCTCCGACATCTTTGACTGTGCTGCTACCTGACCAAGCTC 420
Db |||
Qy 3988 GGCAGCTGGACCACTGGCCCTCCGACATCTTTGACTGTGCTGCTACCTGACCAAGCTC 4047
Db |||
Qy 421 TCCCGGGGTGGACCTTACACCTTCCGCAAGCATGTGTACAAAGGAGGAATGGGTCCC 480
Db |||
Qy 4048 TCCCGGGGTGGACCTTACACCTTCCGCAAGCATGTGTACAAAGGAGGAATGGGTCCC 4107
Db |||
Qy 481 TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCACCTGGGCCCTCTGAGGAG 540
Db |||
Qy 4108 TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCACCTGGGCCCTCTGAGGAG 4167
Db |||
Qy 541 GAGAGCCAGGGGGCGGTACGCCCAACCCCTGCCAGCACAAAGACCTTGGCAATCCAGACA 600
Db |||
Qy 4168 GAGAGCCAGGGGGCGGTACGCCCAACCCCTGCCAGCACAAAGACCTTGGCAATCCAGACA 4227
Db |||
Qy 601 CAGATCCAGAGGGGCGGTTCAGCGTGGTGGGCAATGCTGGGAGAGGCCAGCGGGCGG 660
Db |||
Qy 4228 CAGATCCAGAGGGGCGGTTCAGCGTGGTGGGCAATGCTGGGAGAGGCCAGCGGGCGG 4287
Db |||
Qy 661 GCGTGGCCGCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGCAGTGTCTGCGGAA 720
Db |||
Qy 4288 GCGTGGCCGCAAGATCATCCCTTACCAACCCCAAGGACAAAGACAGCAGTGTCTGCGGAA 4347
Db |||
Qy 721 TAGAGGCCCTCAAGGGCGTGGCCACCGCACCTGGCCCGAGCTGCAAGCAGCTTACCTC 780
Db |||
Qy 4348 TAGAGGCCCTCAAGGGCGTGGCCACCGCACCTGGCCCGAGCTGCAAGCAGCTTACCTC 4407
Db |||
Qy 781 AGCCCCGGCACCTGGTGTCTCATCTTGAGAGTGTGCTCTGGGGCCCGAGCTGCTCCCTGC 840
Db |||
Qy 4408 AGCCCCGGCACCTGGTGTCTCATCTTGAGAGTGTGCTCTGGGGCCCGAGCTGCTCCCTGC 4467
Db |||
Qy 841 CTGGCCGAGAGGGCCCTCCTACTCAGAAATCCGAGGTGAAGGACTACCTGTGCGAGATGTTG 900
Db |||
Qy 4468 CTGGCCGAGAGGGCCCTCCTACTCAGAAATCCGAGGTGAAGGACTACCTGTGCGAGATGTTG 4527
Db |||
Qy 901 AGTGCCACCCAGTACTCGCAACCCAGCACATCTTGACACTGGACCTGAGGTCCGAGAAC 960
Db |||
Qy 4528 AGTGCCACCCAGTACTCGCAACCCAGCACATCTTGACACTGGACCTGAGGTCCGAGAAC 4587
Db |||
Qy 961 ATGATCATCACCGAATACAACCTGCTCAAGTCTGTTGACCT 1001
Db |||
Qy 4588 ATGATCATCACCGAATACAACCTGCTCAAGTCTGTTGACCT 4628
Db |||
```

Search completed: March 21, 2005, 16:07:45
Job time : 633.052 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 16:50:48 ; Search time 3663.47 Seconds
(without alignment)
10400.615 Million cell updates/sec

Title: US-10-077-130-4_COPY_3120_4120
Perfect score: 1001
Sequence: 1 aagatgggaagagctgag.....tcagagcccaaggcgtgtt 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	428.2	42.8	553	CN302349	CN302349 170005326
2	424.2	42.4	621	CF179218	CF179218 813316 MA
3	422.6	42.2	632	CF180572	CF180572 815696 MA
4	377.6	37.7	538	CN302346	CN302346 170005326
5	254.6	25.4	591	BU439885	BU439885 604144795
6	215.4	21.5	1142	BG680164	BG680164 602626828
7	212.2	21.2	401	AW898791	AW898791 CM0-NN007
8	208	20.8	786	CD099595	CD099595 AGENCOURT
9	207.6	20.7	522	AQ595824	AQ595824 HS 2132 B
10	194.8	19.4	658	AZ383233	AZ383233 1M0140L17
11	193.8	19.4	236	CV426951	CV426951 RC6-EN008
12	193.8	19.4	574	A1595154	A1595154 ml39h10.Y
13	161.2	16.1	588	BU436399	BU436399 604145537
14	148.8	14.9	343	BE159678	BE159678 MR0-HT040
15	140.8	14.1	4130	AK048144	AK048144 Mus muscu
16	140.8	14.1	5558	AK029863	AK029863 Mus muscu
17	140.6	14.0	373	AA061264	AA061264 ml39h10.Y
18	139.8	14.0	631	CG793386	CG793386 UTSW_SMI1
19	132.2	13.2	397	BZ837763	BZ837763 CH240_248
20	127.8	12.8	1192	CC247513	CC247513 CH261-127
21	123.2	12.3	584	CK830169	CK830169 Fr Rev_03
22	123.2	12.3	2066	AV399870	AV399870 Homo sapi
23	123.2	12.3	2274	CR590353	CR590353 full-leng
24	120.4	12.0	472	CE495069	CE495069 tigr-gss-

25	118.2	11.8	179	4	BG989628	BG989628
26	113.2	11.3	980	5	EX343487	EX343487
27	108.6	10.8	770	4	BG819323	BG819323
28	107.8	10.8	2066	9	AY399872	Mus muscu
29	106.2	10.6	1085	4	BM546177	BM546177 AGENCOURT
30	104	10.4	724	1	AU141131	AU141131
31	102.8	10.3	569	4	BM254228	BM254228
32	102.4	10.2	655	1	AU119815	AU119815
33	102.2	10.2	949	5	BQ645083	AGENCOURT
34	101	10.1	612	5	BU947303	1046c03.Y
35	100.4	10.0	886	3	CR607106	full-leng
36	100.4	10.0	886	5	BU538739	AGENCOURT
37	100.4	10.0	886	5	EX377136	EX377136
38	99	9.9	796	5	EX381923	EX381923
39	98.8	9.9	205	7	CV419760	RC4-CN018
40	98.2	9.8	1009	5	BQ653435	AGENCOURT
41	98	9.8	1098	4	BM543727	BM543727
42	97.6	9.8	575	5	BU072543	1m44e07.Y
43	97	9.7	630	7	CN344412	CN344412
44	96.6	9.7	626	2	AW743136	UP63a04.Y
45	96.6	9.7	960	5	BQ645277	AGENCOURT

ALIGNMENTS

RESULT 1
CN302349 553 bp mRNA linear EST 16-MAY-2004
LOCUS 17000532627462 GRN_ES Homo sapiens CDNA 5', mRNA sequence.
DEFINITION CN302349
ACCESSION CN302349
VERSION CN302349.1 GI:47318763
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 553)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Flsek, G. J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M. S., Mandalam, R., Lebkowski, J. and Stanton, L. W.
TITLE Transcriptome characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 553 Std Error: 0.00.
FEATURES
source
1..553
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"
/clone_lib="GRN ES"
/note="Toligo dt-primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Query Match 42.8%; Score 428.2; DB 7; Length 553;
Best Local Similarity 85.9%; Pred No. 2e-89;
Matches 475; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 338 GGAGCTGTGTCTGCACAGCGGGCAAGCAGATGCTGGGAGTACAGCTGTGTGAGGTG 397
DB 1 GAAGCTGTGTCTGCACAGCTGGGCAAGCAGATGCTGGGAGTACAGCTGCGAGGTG 60

Qy	398	GGGGCCAGAGAGTCTCTTCCACCTGCGCATCACAGAGCCCAAGGGGGTGTTCGCAAGG	457
Db	61	GGGGCCAGAGAGTCTCTTTCNACTGCGCATCAGAGAGCCCAAGGCGAGTGTTCGCAAGG	120
Qy	458	AGCAGTCAGTCGCAATAATGAGGTGCGAGGCTGAGGCGGGGACCACTGCCATGCTGAGCTGTG	517
Db	121	AGCAGTTGGTGCATATGATGAGGTGCGGACTGAGGCGAGGGCCAGTGCCACACTGAGCTGTG	180
Qy	518	AGGTGGCCAGGCCAGCCAGACGAGGTGACGTGGTACAGGACGGGAAGAAGCTGAGTCCCA	577
Db	181	AGGTGGCCAGGCCAGCCAGACAGAGGTGAGCTGGTACAGGATGGGAAGAAGCTGAGTCCCA	240
Qy	578	GCTCAAAAGTACGCGATGGAGGTCAAGGCGCTGCACACGAAGGCTGTTAGTCACACAGCTGG	637
Db	241	GTTTCGAAAGTGCCTAGATAGAGGCCGCGGGCTGCTATCGCGCAGCTGGTGTGCACAGCGAG	300
Qy	638	GCAAGCAGATGCTGGGAGTACAGCTGCGAGGCTGGGGGCCAGAGAGTCTCTCTTTCAAC	697
Db	301	GCCAGCAGATGCTGGGAGTACACTGTGAGGCGCGGGGCCAGCGGCTCTCTCTCCACC	360
Qy	698	TGCACATCACAGAGCCCAAGGCGAGTGTGTTGCCAAGGAGCAGTTGGTGCATATATGAGGTGC	757
Db	361	TGGATGTTTCAGAGCCCAAGGGGTGTTTGCNAAGGAGCAGCTGGGCACACAGGAAGGTGC	420
Qy	758	GGACTGAGCGCAGGGGCCAGTGCGCACATGAGCTGTGAGCTGTGGCCCCAGGCCACAGAGG	817
Db	421	AGGCCGAGCGGGGGCCATTGCCGCTGAGCTGCGAGGTGGGCCAGGCCCCAGACAGAGG	480
Qy	818	TGACGTGGTACAAAGGATGGGAAGAAGCTGAGTCTCCAGTTCGAAGGTGGCCTAGAGGGCTG	877
Db	481	TGACGTGGTACAAAGACGGGGAAGAAGCTGAGTCTCCAGTTCGAAGTTCGAATGGAGGGCTG	540
Qy	878	CGGGCTGCATGCG	890
Db	541	TGGGCTGCACAGC	553

RESULT 2	CFI79218	CFI79218	621 bp	mRNA	linear	EST 28-JUL-2003
LOCUS						
DEFINITION			813316	MARC 3P1G	Sus scrofa	cdna 5', mRNA sequence.

EST.	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
EST. 922226.1						
		Sus scrofa (pig)				
		Sus scrofa				
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.				
		1 (bases 1 to 621)				
		1				
					Smath, T.P.L., Freking, B.A., Ford, J.J., Vallet, J.L., Fox, J., Wise, T.A., Nonneman, D.J., Wray, J.E. and Keele, J.W.	
						A second set of porcine ESTs from a pooled-tissue normalized

Unpublished (2003)
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: SRG8009 row: 0 column: 4
Seq primer: GTAATACGACTCACTATAGG.

/notes="Vector: pCDNA3.1; Site 1: EcoRI; Site 2: NotI; Library made with RNA pooled from multiple tissues including brain, liver, muscle, placenta/embryum, ovary, testes, and bone marrow."

	Query Match	42.4%;	Score 424.2;	DB 7;	Length 621;	
	Best Local Similarity	80.2%;	Pred. No. 1.8e-88;			
	Matches 498;	Conservative	0;	Mismatches 123;	Indels	Gaps 0;
QY	108	CAGCTGCGAGGCCAGGGGCCACAGGGTCTCTTCGGCTGCACATCA	CAGAGCCCAAGAT	167		
Db	1	CAGCTGCGAGGTGGGGGCCAAAAAGTCTCTTCACCTGGATGTCA	CATAGCCTCAGT	60		
QY	168	GATGTTTCAAAGGACGAGTCAGTGCCATAATAGGTGCAGGCTCAG	GGCGGGCCAGTGC	227		
Db	61	GGTGTTTTCCAAAGAACAGCTGGCTGCGAGTGAAGTGC	AAGCCATGGCAGGGGCCATTGC	120		
QY	228	CATGCTGAGCTGTGAGGTGGGCCACAGCCCGACGAGGTGACGTGG	TACTAAGATGGAA	287		
Db	121	CACGTGAGCTGTGAGGTGGGCCATGCCCATACGGAGGTGACTTGG	TACTAAGGATGGAA	180		
QY	288	GAAGCTGAGCTCCAAGCTCAAAAGTGGGCATGGAGGTCAAAAGGT	GCACACGAGGCTGGT	347		
Db	181	GAAGCTGAGTTGCGAGCTCAAAAGTGC GTGTGGAGGCCAAGGGCT	GC AACCTGGAGGCTGGT	240		
QY	348	GCTGCCACAGCGGGSCAAAGCAGATGCTGGGAGTACAGCTGTGAGG	CTGGGGCCACAG	407		
Db	241	GGTGATCAAGTGGGCCAAGGCATATGCTGGGAGTACAGCTGCGATG	CTGGGGGTCAAA	300		
QY	408	AGTCTCTTCCACCTGCACATCA CAGAGCCC AAGGGGTGT TTTGCGAAG	GAGCAGTCAGT	467		
Db	301	GGTCTCTTCCACCTGGATGTCA CATATCCCTCAGTGTGT TTTGCCAAG	GAGCAGCCATC	360		
QY	468	GCATAATGAGGTGACGCTGAGCGGGGACA CTGCCATGCTGAGCTGT	GAGTGGCCCCA	527		
Db	361	ACGCACCTGAAGTGC AAAGCCATGGCCGAGGCAGTGC CCACACTG	AGCTGTGAGGTGGCCCCA	420		
QY	528	GCCCCAGACGGAGGTGACGTGGTAC AAGGACGGGAAGAAGCTCAGCT	CACAGCTCAAAGT	587		
Db	421	TGCCACACAGAGGTGACATGGTACAAGGATGGGAGAGCTGACTTCA	AGCTCAAAGT	480		
QY	588	ACGCATGGAGGTCAAGGGCTGCACAG AAGGCTGGTAGTGCAGCAGGT	GGGGCAAGCAGA	647		
Db	481	CGGTGTGGAGGCCAATGGCTGCAACCGGATGCTGGTGGTGCATCATG	CGGGCAATGCATA	540		
QY	648	TGCTGGGGAGTACAGCTGCGAGCTGGGGGCCACAGAGTCTCTTTCACTG	TCACACATCAC	707		
Db	541	TGCTGGGGAGTACAGCTGCGATGCTGGGGGTCAATATGCTCTCTTCC	ACTGGACATCAC	600		
QY	708	AGAGCCCAAGCAGTGT TTGCG	728			
Db	601	AGAGCCCTCAGTGGTGT TTGCG	621			

[illegible]

USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390

Email: smiths@mail.marc.usda.gov
Single pass sequencing. Bases called with phred v0.020425.c and
trimmed with the aid of the trim_alt option. Vector identified with
cross_match v0.990329.
Plate: SR88009 row: O column: 4
Seq primer: TAGAAGGCACATCTCAGG.

Location/Qualifiers

FEATURES
source

1. .632
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 3P1G"
/note="Vector: pcDNA3.1; Site 1: EcoRI; Site 2: NotI;
Library made with RNA pooled from multiple tissues
including brain, liver, muscle, placenta/endometrium,
ovary, testes, and bone marrow."

ORIGIN

Query Match 42.2%; Score 422.6; DB 7; Length 632;

Best Local Similarity 79.5%; Pred. No. 4.2e-88; Mismatches 129; Indels 0; Gaps 0;

Matches 500; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 147 GCACATCACAGAGCCCAAGATGATGTTTCCAAAGGAGCAGTCAGTGCATATGAGGTGCA 206
DB 632 GGATGTCACAGAGCCCTCAGTGGTGTGTTGCCAAGNACAGCTGCGCTGCAGTGAAGTGA 573
QY 207 GGCTAGGCGGGGGCAGTGCCATGCTGAGCTGTGAGGTGGCCAGGCCACAGCGAGGT 266
DB 572 AGCCATGCGAGGGGCGAGTGCACAGTGTGAGGTGGCCAGGCCACAGCGAGGT 513
QY 267 GAGTGTGTACAGAGTGGGAGAGCTGAGCTCCAGCTCAGAGTCCAAAGTGGGCGAGGTCAA 326
DB 512 GACTTGGTACAGAGTGGGAGAGCTGAGTTGAGCTCAAAGTGGCGTGGAGGCCAA 453
QY 327 AGGTGTCACAGAGGCTGGTGTCTGCCACAGCGGGGCAAGCAGATGCTGGGAGTACAG 386
DB 452 GGGCTGCAACCGAGGCTGGTGTGCAGCAAGTGGGCAAGCAGATGATGGGAGTACAG 393
QY 387 CTGTAGGCTGGGGGCGAGAGTCTCTTCCACCTGCACATCAGAGGCCCAAGGGGT 446
DB 392 CTGCGAGGCTGGGGGTCAAGAGTCTCTTCCACATGATGTACAGAGCCCTCAGTGT 333
QY 447 GTTTCGAGGAGCAGTCAAGTGCATATGAGGTGAGGTGAGGCGGGGACCTGCCAT 506
DB 332 GTTTCGAGGAGCAGCAGCAGCAGCAATGAAGTGAAGCCATGGCGGAGCCAGTGGCCAC 273
QY 507 GCTGAGCTGTGAGGTGGGCGGAGGAGGTGAGCTGGTACAAAGGAGCGGAAGAA 566
DB 272 AATGAGCTGTGAGGTGGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 213
QY 567 GCTGAGCTCCAGCTCAAAGTACGATGAGAGTCAAGGCTGCAACAGAGGCTGGTGT 626
DB 212 GCTGACTTCAAGCTCAAAGTGCAGTGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 153
QY 627 GCAGAGGCTGGGCAAGCAGATGCTGGGAGTACAGCTGCGAGGCTGGGCGGCGGAGGT 686
DB 152 GCAGAGGCGGCGGCAAGGAGGAGTGTGGGAGTACAGCTGCGAGGCTGGGCGGCGGAGGT 93
QY 687 CTCCTTTCAACTGCATCACAGAGCCCAAGGAGTGTGTCAGAGGAGCAGTGGTGA 746
DB 92 CTCCTTCCATGAGATCAAGAGCCCTCAGTGTGTTGTCAGAGGAGCAGCCATCAG 33
QY 747 TAATGAGGTGGGAGCTGAGGCGAGGGGCA 775
DB 32 CAGTCCAGTGGAGGTGTGACGGGGGCA 4

RESULT 4

CH302346

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

1. .538

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"

/clone_lib="GRN ES"

/note="oligo dT primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN

Query Match 37.7%; Score 377.6; DB 7; Length 538;

Best Local Similarity 84.3%; Pred. No. 1.3e-77;

Matches 425; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 482 AGGCTGAGCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCAGCCCGCAGCGAGG 541

DB 1 AGGCTGAGCGCGGGCCAGTCCACATGAGCTGTGAGGTGGCCAGCGAGG 60

QY 542 TGAGCTGTGTACAGGACGGAAGCTGAGCTCCAGCTCAAAAGTACGATGGAGGTCA 601

DB 61 TGACGTGTGTACAGGATGGGAAGAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGCGCG 120

QY 602 AGGGCTGCACAGAGGTGGTGTGTGACAGCAGGTGGGCAAGCAGATGCTGGGGAGTACA 661

DB 121 CGGGCTGCATCGGCGAGCTGGTGTGACAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 180

QY 662 GCTGCGAGGCTGGGGGCGAGAGTCTCTTTCATCTGCACATCAGAGCCCAAGGCGAG 721

DB 181 CTTGTGAGCGGGGGCGAGCGGCTCTCTTCCACCTGATGTTTCAAAGCCCAAGGCGG 240

QY 722 TGTGTTCCAAAGGAGCAGTGGTGTGCAATATGAGTGGGAGTGGAGGCGGCGGCGGCGG 781

DB 241 TGTGTTCCAAAGGAGCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300

QY 782 CACTGAGCTGTGAGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 841

DB 301 CGCTGAGCTGCGAGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360

QY 842 AGCTGAGCTCCAGTTTCAAAGTGGCGATGAGAGGCTGCGGGCTGCATGCGGCGAGCTGGTGG 901

DB 361 AGCTGAGCTCCAGCTCGAAAGTTCGAAATGAGGCTGTGGGCTGCACACGAGGCTGGTGG 420

QY 902 TGCAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 961

CH302346 538 bp mRNA linear EST 16-MAY-2004
1700532600207 GRN_ES Homo sapiens cDNA 5', mRNA sequence.
CH302346
EST. GI:47318760

Homo sapiens (human)
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Flak, G. J.,
Li, Y., Xu, C., Pang, R., Guebler, K., Rao, M. S., Mandalam, R.,
Lebkowski, J. and Stanton, L. W.
Transcriptome characterization elucidates signaling networks that
control human ES cell growth and differentiation
Nat. Biotechnol. 22 (6), 707-716 (2004)

Contact: Brandenberger R
Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com

Insert Length: 538 Std Error: 0.00.
Location/Qualifiers
1. .538
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="embryonic stem cells, cell lines H1, H7, and H9"

/clone_lib="GRN ES"
/note="oligo dT primed, full-length enriched cDNA library from undifferentiated hES cell lines H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free conditions"

ORIGIN
Query Match 37.7%; Score 377.6; DB 7; Length 538;
Best Local Similarity 84.3%; Pred. No. 1.3e-77;
Matches 425; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

QY 482 AGGCTGAGCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCAGCCCGCAGCGAGG 541
DB 1 AGGCTGAGCGCGGGCCAGTCCACATGAGCTGTGAGGTGGCCAGCGAGG 60

QY 542 TGAGCTGTGTACAGGACGGAAGCTGAGCTCCAGCTCAAAAGTACGATGGAGGTCA 601
DB 61 TGACGTGTGTACAGGATGGGAAGAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGCGCG 120

QY 602 AGGGCTGCACAGAGGTGGTGTGTGACAGCAGGTGGGCAAGCAGATGCTGGGGAGTACA 661
DB 121 CGGGCTGCATCGGCGAGCTGGTGTGACAGAGGCGGCGGCGGCGGCGGCGGCGGCGG 180

QY 662 GCTGCGAGGCTGGGGGCGAGAGTCTCTTTCATCTGCACATCAGAGCCCAAGGCGAG 721
DB 181 CTTGTGAGCGGGGGCGAGCGGCTCTCTTCCACCTGATGTTTCAAAGCCCAAGGCGG 240

QY 722 TGTGTTCCAAAGGAGCAGTGGTGTGCAATATGAGTGGGAGTGGAGGCGGCGGCGGCGG 781
DB 241 TGTGTTCCAAAGGAGCAGTGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 300

QY 782 CACTGAGCTGTGAGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 841
DB 301 CGCTGAGCTGCGAGGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 360

QY 842 AGCTGAGCTCCAGTTTCAAAGTGGCGATGAGAGGCTGCGGGCTGCATGCGGCGAGCTGGTGG 901
DB 361 AGCTGAGCTCCAGCTCGAAAGTTCGAAATGAGGCTGTGGGCTGCACACGAGGCTGGTGG 420

QY 902 TGCAGCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 961

Db 421 TCCAGCAGGATGCCAGGCGGACACCGGGAGTATAGCTGCGAGGCGGGGCCAGCGGC 480

QY 962 TCTCCTTCCACCTGGATGTTTCAG 985

Db 481 TCTCCTTCCAGCTGGACGTGGCAG 504

RESULT 5
BU439885
LOCUS 591 bp mRNA linear EST 29-NOV-2002
DEFINITION 604144795F1 CSEQRN11 Gallus gallus cDNA clone ChEST983m23 5', mRNA
sequence.
ACCESSION BU439885
VERSION BU439885.1 GI:25929196
KEYWORDS EST.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
1 (bases 1 to 591)
Boardman, P.E., Sanz-Ezquerro, J., Overton, I.M., Burt, D.W., Bosch, E.,
Fong, W.T., Tickle, C., Brown, W.R.A., Wilson, S.A. and Hubbard, S.J.
A Comprehensive Collection of Chicken cDNAs
Curr. Biol. 12 (22), 1965-1969 (2002)
22335534
12445392
Contact: Simon Hubbard
Department of Biomolecular Sciences
University of Manchester Institute of Science and Technology
(UMIST)
PO Box 88, Manchester, M60 1QD, UK
Tel: 01612008930
Fax: 01612360409
Email: Simon.Hubbard@umist.ac.uk.

FEATURES

source
1..591
Location/Qualifiers
/organism="Gallus gallus"
/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone="ChEST983m23"
/sex="Male and female"
/tissue_type="muscle"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CSEQRN11"
/note="Vector: pBluescript II KS(+); Site 1: EcoRI;
Site 2: NotI; This normalized library was constructed from
1 million independent clones. cDNA synthesis was initiated
using an oligo(dT) primer, using methylated C in the first
strand synthesis reaction. Following this first strand
reaction, double-stranded cDNA was blunt-ended, ligated to
NotI adapters, digested with EcoRI, size-selected, and
cloned into the NotI and EcoRI compatible sites of a
custom modified MCS of the pBluescript (KS+) vector. The
library was normalized in 2 rounds using conditions
adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
Bonaldo et al., Genome Research 6 (1996): 791, except that
a significantly longer reannealing hybridization was
used."

ORIGIN

Query Match 25.4%; Score 254.6; DB 5; Length 591;
Best Local Similarity 65.7%; Pred. No. 7.8e-49;
Matches 371; Conservative 0; Mismatches 194; Indels 0; Gaps 0;
431 CAGAGCCACAGGGGTTTTCGAGGAGGAGCAGTGCATGTCATATGAGTGCTGAGG 490
1 CAGAGCCAGAGTGTGTTATTCATCAACAGGAGAGAGGTGCAGAGAGGTGAAGCTGCAC 60
491 CGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCAGCCCGAGCGAGGTGACGTGGT 550

Db 61 CAACAGAAAATGCCACGCTGAGCTGCGAGGTGGCGAGGAGACGAGCGAGTGAAGTGGT 120

QY 551 ACAAGACGGGAAGAAGCTGAGCTCCAGCTCAAAAGTACGATGAGAGTCAAGGGCTGCA 610

Db 121 ACAAGGAGGGGAAAACCTGATCACTCGAGCAAGAAGTTTCAGGGTGGAGTTCAGAGGGCAAAAC 180

QY 611 CACGAAGGCTGGTAGTCAGCAGGTGGGCAAGCAGATGCTGGGGAGTACAGCTGCCGAGG 670

Db 181 TGGCTCGGCTGGTGGTGGCCAGGTGGAGAAGAAGATGCAGGGGAGTACACCTGCCGAGG 240

QY 671 CTGGGGGCCAGAGAGTCTCTTTCAACTGCACATCAGAGAGCCCAAGGCACTGTTTGCAC 730

Db 241 CTGCTGCCAGAAACTGACCTTCAAGATTGATGTACAGAGCCAAACCTGTGTTTATAA 300

QY 731 AGGAGCAGTGGTGCATATGAGGTGCGGATGAGGAGGCGGCGCCAGTGCACACTGAGCT 790

Db 301 ACCAGGAGAGGTGCAGAGGGAGGTGAGTGTCTCTGGCAGAAAAGTCCGCCCTCAGCT 360

QY 791 GTGAGGTGGCCCGCCAGCCAGACAGAGTGCCTGGTACAAGGATGGGAAGAAGCTGAGCT 850

Db 361 GTGAAGTAGCACAGATACACCGAGTAAATGGTACAGGATGGAAGACTGCTTGCCT 420

QY 851 CCAGTTCGAAGTCCGATAGAGGCTCGGGCTGCATGCGGACGCTGGTGTGAGCAGG 910

Db 421 CCTCTAGAAAAGTTCAAAATGGAACCTGTGGGCAAAACCCGCGCTCTGGTTGTGAGCAGC 480

QY 911 CAGGCCAGGAGTGTGGGGAGTACACCTGTGAGGCTGGGGCGCAGGCTCTCCTTCC 970

Db 481 TGGAGAAGAAGATGCTGGAGAATACTGTGTGAGGCTGCGCGCAGGAGCTGACCTTCA 540

QY 971 ACCTGGATGTTTCAGAGCCCAAGGC 995

Db 541 AGTTGGAAGCACTGAACAGAGGC 565

RESULT 6
BG680164/c
LOCUS 1142 bp mRNA linear EST 01-MAY-2001
DEFINITION 602626828F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4751671 5',
mRNA sequence.
BG680164
BG680164.1 GI:13911561
EST.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1142)
NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAMI0608 row: h column: 08
High quality sequence stop: 708.
Location/Qualifiers
1..1142
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4751671"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI CGAP Skn4"
/note="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.

FEATURES

source
1..1142
Location/Qualifiers

Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI-CGAP Library."

```

ORIGIN
Query Match      21.5%; Score 215.4; DB 4; Length 1142;
Best Local Similarity 84.1%; Pred. No. 1.3e-39;
Matches 243; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

Qy 697 CTGCACATCAGAGCCCAAGCGAGTGTTCACAGGAGCAGTGTGTCATAATGAGGTG 756
    |||||
Db 322 CTGCTCTCCAGAGCCCAAGCGGTGTTCAGGAGCAGCCAGCAGCAGAGGAGGTG 263
    |||||

Qy 757 CGGACTGAGGCGGGCCAGTGCACACTGAGCTGAGCTGAGGTGGCCAGCCAGCAGAG 816
    |||||
Db 262 CAGGCTGAAGCAGGGGCTAGGCCACACTGAGCTGAGGTGGCCAGCCAGCAGAG 203
    |||||

Qy 817 GTGACGTGTGTAACAGGATGGGAAGCTGAGCTCCAGTTTGAAAGTGCATAGAGGCT 876
    |||||
Db 202 GTGACTTGTGTACAGGACGGGAAGAGTGTGAGCTCCAGTTGAAAGTGTGAGGCT 143
    |||||

Qy 877 GCGGGCTGCATGCGGCAGCTGTGTGTCAGCAGGAGCCAGCCAGCTGTGGGGAGTAC 936
    |||||
Db 142 GCAGGCTGTACAGGAGCTGTGTGTCAGCAGGAGCCAGCCAGCAGCAGAGGATAT 83
    |||||

Qy 937 ACCTGTGAGGCTGGGGCCAGCGCTCTCCTTCCACCTGGATGTTTCAG 985
    |||||
Db 82 AGCTGTGAGGCTGGGGCCAGCAGCTCTCCTTCCGCTGCAAGTGGCAG 34
    |||||

```

```

RESULT 7
AW898791
LOCUS      AW898791
DEFINITION CMO-NN0075-130400-332-f06 NN0075 Homo sapiens cdna, mRNA sequence.
ACCESSION AW898791
VERSION    AW898791.1 GI:8062996
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens

```

```

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 401)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H.,
Brunstein, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
10737800

```

```

CONTACT: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=4t2=CMO-NN0075-130
400-332-f06&t3=2000-04-13&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 401.
Location/Qualifiers
1..401
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN0075"

```

```

FEATURES
source
1..401
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="NN0075"

```

/note="Organ: nervous normal; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

ORIGIN
Query Match      21.2%; Score 212.2; DB 2; Length 401;
Best Local Similarity 83.4%; Pred. No. 6e-39;
Matches 241; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

Qy 697 CTGCACATCAGAGCCCAAGCGAGTGTTCACAGGAGCAGTGTGTCATAATGAGGTG 756
    |||||
Db 61 CTGCTCTCCAGAGCCCAAGCGGTGTTCACAGGAGCAGCCAGCAGGAGGAGGTG 120
    |||||

Qy 757 CGGACTGAGGCGGGCCAGTGCACACTGAGCTGTGAGTGGCCAGCCAGCCAGCAGAG 816
    |||||
Db 121 CAGCTGAGGCGGGGAGCAGTGCACGCTGAGCTGCGAGGTGGCCAGCCAGCAGAG 180
    |||||

Qy 817 GTGACGTGTGTAACAGGATGGGAAGCTGAGCTCCAGTTTGAAAGTGCATAGAGGCT 876
    |||||
Db 181 GTGACGTGTGTAACAGGACGGGAAGAGTGTGAGCTCCAGTCCGAAAGTGCAGAGGCC 240
    |||||

Qy 877 GCGGGCTGCATGCGGCAGCTGTGTGTCAGCAGGAGCCAGCCAGCTGTGGGGAGTAC 936
    |||||
Db 241 GTGGGCTGCACACGAGGCTGTGTGTCAGGAGCCAGCCAGCCAGCCGCGGGAGTAC 300
    |||||

Qy 937 ACCTGTGAGGCTGGGGCCAGCGCTCTCCTTCCACCTGGATGTTTCAG 985
    |||||
Db 301 AGCTGTGAGGCGGGGATCAGCGCTGTCTTCCAGCTGCAGTGGCAG 349
    |||||

```

```

RESULT 8
CD099595
LOCUS      CD099595
DEFINITION AGENCOURT_13997453 NICHD_XGC_Tad1 Xenopus laevis cdna clone
IMAGE:6940598 5', mRNA sequence.
ACCESSION CD099595
VERSION    CD099595.1 GI:30752698
KEYWORDS   EST.
SOURCE     Xenopus laevis (African clawed frog)
ORGANISM   Xenopus laevis

```

```

REFERENCE
1 (bases 1 to 786)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Drs. Donald Brown and Lihuan Cai
cDNA Library Preparation: CLONTECH
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUCM3267 row: e column: 13
High quality sequence stop: 437.
Location/Qualifiers
1..786
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6940598"
/dev_stage="metamorphosis stage 53"
/clone_lib="NICHD_XGC_Tad1"
/note="Organ: Developing Tadpole; Vector: pDNR-LIB;
Site_1: Sfi; Site_2: Sfi; 5' and 3' adaptors were used in

```


84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunnett@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0140 row: L column: 17
Seq primer: CACACAGGAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 658.
Location/Qualifiers
FEATURES
source
1..658
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0140L17"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, P-"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/note="Vector: pMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 236)
AUTHORS Dias Neto,E., Garcia Correa,R., Varjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,P.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. http://www.ludwig.org.br.
FEATURES
source
1..236
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="EN0083"
/note="Organ: lung_normal; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196.716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN
Query Match 19.4%; Score 193.8; DB 7; Length 236;
Best Local Similarity 96.6%; Pred. No. 1,1e-34;
Matches 198; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 664 TCGCAGGCTGGGGCCAGAGAGTCTCTTTCAACTGCACATCACAGAGCCCAAGCAGTG 723
DB 32 TGCRAAGTTGCAGGCCAGAGAGTCTCTTTCAACTGCACATCACAGAGCCCAAGCAGTG 91
QY 724 TTTCGAAGGACAGTCTGGTGCATTAATGAGTCCGAGCTGAGGCGAGGCGGCGAGTCCACA 783
DB 92 TTTCGAAGGACAGTCTGGTGCATTAATGAGTCCGAGCTGAGGCGAGGCGGCGAGTCCACA 151
QY 784 CTGAGCTGTAGGTGGCCCGAGCCAGACAGAGGTGAGTGGTACAAAGATGGGAAGAG 843
DB 152 CTGAGCTGTAGGTGGCCCGAGCCAGACAGAGGTGAGTGGTACAAAGATGGGAAGAG 211
QY 844 CTGAGCTCCAGTTCGAAAGTGCACA 868
DB 212 CTGAGCTCCAGTTCGAAAGTGCACA 236
RESULT 12
LOCUS AI595154/c
DEFINITION ml39h10.v1 Stratagene mouse testis (#937308) Mus musculus cDNA
ACCESSION AI595154
VERSION AI595154.1 GI:4604202
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Query Match 19.5%; Score 194.8; DB 8; Length 658;
Best Local Similarity 81.3%; Pred. No. 7.6e-35;
Matches 226; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 156 AGAGCCCAAGATGATGTTTGCACAGAGCAGTCACTGCATATAGTGCAGGTCAGGCGAGGC 215
DB 182 AGACACCAAGGCTGATGTTTGCACAGAGCAGCAGCAGCCAGGCTGAAGGCTGAGGC 241
QY 216 GGGGCCCATGCTGAGTGTGAGTGGCCCGAGCCAGCGAGGTGACCTGCTA 275
DB 242 GGGGAACAGTGCACCCCTGAGCTGGAGGTGGCCCGAGCCAGACTGAGGTGACATGTT 301
QY 276 CAAGGATGGGAAGAGCTGAGTCTCAGCTCAAAAGTGGGCATGGAGGTCAAAAGGCTGCAC 335
DB 302 CAAGACGGGAAGAGCTGAGTCTCAGCTCGAAGGTGCGGTGGAGGCTCGGGCTGCTC 361
QY 336 ACGGAGGCTGTGCTGCCACAGCGCGGCAAGCAGATGCTGGGGAGTACAGCTGTGAGGC 395
DB 362 CAGGAGGCTGTGGTGCAGCAGCGCGGCAAGCGGATGCTGGGGAGTACAGCTGCGAGGC 421
QY 396 TGGGGCCAGAGAGTCTCTTCCACTGCACATCACAG 433
DB 422 CGGGGGTCAGAGAGTCTCTCTCCGCTCGACGTGGCAG 459
RESULT 11
LOCUS CV426951
DEFINITION RC6-EN0083-311000-011-H11 EN0083 Homo sapiens cDNA, mRNA sequence.
ACCESSION CV426951
VERSION CV426951.1 GI:52822454
KEYWORDS EST.
SOURCE Homo sapiens (human)

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 574)
Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T.,
Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y.,
Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R.,
Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R.,
Waterston,R. and Wilson,R.
The WashU-NCI Mouse EST Project 1999

Unpublished (1999)
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LLNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 This read is a RESQUENCE of a previously sequenced mouse clone
 This read has been verified (found to hit its original self in the
 correct orientation)

FEATURES	Location/Qualifiers
source	1. .574

FEATURES
source

```
1. 588
/organism="Gallus gallus"
/mol_type="mRNA"
```

```

/mol_type="mRNA"
/strain="Layer and broiler"
/db_xref="taxon:9031"
/clone="ChST98519"
/sex="Male and female"
/tissue_type="muscle"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="CS98RN11"
/note="Vector: pBluescript II KS(+); Site 1: EcoRI;
Site 2: NotI; This normalized library was constructed from
1 million independent clones. cDNA synthesis was initiated
using an oligo(dT) primer, using methylated C in the first
strand synthesis reaction. Following this first strand
reaction, double-stranded cDNA was blunt-ended, ligated to
NotI adapters, digested with EcoRI, size-selected, and
cloned into the NotI and EcoRI compatible sites of a
custom modified MCS of the pBluescript (KS+) vector. The
library was normalized in 2 rounds using conditions
adapted from Soares et al., PNAS (1994) 91: 9228-9232 and
Bonaldo et al., Genome Research 6 (1996): 791, except that
a significantly longer reannealing hybridization was
used."

```

ORIGIN

Query Match 16.1%; Score 161.2; DB 5; Length 588;
Best Local Similarity 67.1%; Pred. No. 5.4e-27;
Matches 243; Conservative 0; Mismatches 118; Indels 1; Gaps 1;

Qy	619	CTGGTAGTCGACGAGTGGGCAAGCAGATGCTGGGGAGTACAGCTGCGAGGCTGGGGGC	678
Db	1	CTGGTGGTGAGCCAGGTGGAGAGAAAGATGC-AGGGAGTACACCTGCGAGGCTGTGGC	59
Qy	679	CAGAGAGTCTCCTTTCAACTGCACATCACAGAGCCCAAGCGAGTGTTTGCCAAGGAGCAG	738
Db	60	CAGAAATCAGCCTTCAAGATTACTGTGCACAGAGCAGAGGATGCTTTTATCAACAGGAG	119
Qy	739	TTGGTGCATAATGAGTGGCGGACTGAGGCAGGGGCCAGTGCACACATGAGCTGTGAGGTG	798
Db	120	AAGTGCAGAGGAGGTGAAGGCTGCACCAACAGAAATGCCACGCTGAGCTGCGAGGTG	179
Qy	799	GCCAGGCCCCACAGACAGAGTGACGTGGTCAAGAGATGGGAAGAGCTGAGCTCCATTTCG	858
Db	180	GGCCAGGAGAAACCGAGGTGAAGTGGTCAAGAGAGGGGAAACTGTATCACCTCGAGCAAG	239
Qy	859	AAAGTCGCATAGAGGCTCGCGGGCTGCATGCGCGCAGCTGGTGGTGCAGCAGGCGAGGCCAG	918

RESULT 13	588 bp	linear	EST 29-NOV-2002
BU436399		mrna	
LOCUS	604145537f1	CSEQREN11 Gallus gallus	CDNA clone CHE98519 5', mRNA
DEFINITION			sequence.

Db 240 AAGTTCCAGGTGGAGTCAGAGGCAAACTGCGTGGCTGGTGGTGGAGCCAGCGTGGAGGAG 299

Qy 919 CGAGATGCTGGGGAGTACACCTGTGAGCTGGGGCCAGCGGCTCTCTTCCACCTGGAT 978

Db 300 AGAAATCGAGGGGAGTACACCTGCGAGGCTGCTGGCCAGAAAGTGAGCTTCAAGATTACT 359

Qy 979 GT 980

Db 360 GT 361

RESULT 14

LOCUS BE159678 343 bp mRNA linear EST 21-JUN-2000

DEFINITION MR0-HT0407-180400-015-e07 HT0407 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE159678

VERSION BE159678.1 GI:8622399

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Dias Neto, E., Garcia Correa, R., Verlovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, P.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bata, G.S., Simpson, D.H., Brunstein, A., Golliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.

TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

MEDLINE 20202663

PUBMED 10737800

COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=et2-MR0-HT0407-180400-015-e07&t3=2000-04-18&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 343.
Location/Qualifiers
1. .343
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0407"
/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORSITES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

FEATURES

source

ORIGIN

Query Match 14.9%; Score 148.8; DB 2; Length 343;
Best Local Similarity 88.3%; Pred. No. 3.9e-24;
Matches 173; Conservative 0; Mismatches 22; Indels 1; Gaps 1;

Qy 153 CACAGAGCCCAAGATGATGTTTGCAGAGGAGTCAGTGCATATGAGTGGCAGGCTGA 212

Db 228 CCCAGAGCCCAAGAGTGGTGTGGCCAGGAGCAGCCACACAGGAGGTGCAGGCTGA 169

Qy 213 GCGCGGGGCCAGTGCCTGCTGAGCTGTGAGGTGGCGCCAGCCAGAGCGGAGGTGACGCTG 272

Db 168 GCGCGGGGCCAGTGCCTGAGCTGCGAGGTGGCGCCAGCCAGAGAGGTGACGCTG 109

Qy 273 GTACAAGGATGGGAAGAGCTGAGCTCCAGCTCAAAAGTGGGCGATGGAGGTCAAAAGGCTG 332

Db 108 GTACAAGGATGGGAAGAGCTGAGTTCAGCTCGAAAGTGGCGCTGGA-GCGCGTGGCTG 50

Qy 333 CACACGGAGGCTGGTG 348

Db 49 CACACGGAGGCTGGTG 34

RESULT 15

LOCUS AK048144

DEFINITION Mus musculus 16 days embryo head cDNA, RIKEN full-length enriched library, clone: C130037H06 product: hypothetical immunoglobulin and major histocompatibility complex domain containing protein, full insert sequence.

ACCESSION AK048144

VERSION AK048144.1 GI:26339169

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS Carninci, P. and Hayashizaki, Y.

TITLE High-efficiency full-length cDNA cloning

JOURNAL Meth. Enzymol. 303, 19-44 (1999)

MEDLINE 92779253

PUBMED 10349636

REFERENCE

AUTHORS Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Muramatsu, M. and Hayashizaki, Y.

TITLE Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

JOURNAL Genome Res. 10 (10), 1617-1630 (2000)

MEDLINE 20499374

PUBMED 11042159

REFERENCE

AUTHORS Shibata, K., Itoh, M., Mizawa, K., Nagao, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

TITLE RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer

JOURNAL Genome Res. 10 (11), 1757-1771 (2000)

MEDLINE 20530913

PUBMED 11076861

REFERENCE

AUTHORS The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

TITLE Functional annotation of a full-length mouse cDNA collection

JOURNAL Nature 409, 685-690 (2001)

REFERENCE

AUTHORS The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE

AUTHORS Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsumi, N., Okazaki, Y., Saito, R., Saitoh, K., Sakai, C., Sakai, K., Sakazume, N.,

Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shinagawa,A., Shiraki,T., Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S., Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A., Muramatsu,M. and Hayashizaki,Y.

TITLE
JOURNAL

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.

FEATURES

Source

1. .4130
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C130037H06"
/db_xref="taxon:10090"
/clone="C130037H06"
/tissue_type="head"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="16 days embryo"
660. .3491
/note="unnamed protein product; hypothetical
immunoglobulin and major histocompatibility complex domain
containing protein (InterPro|IPR003006, evidence:
InterPro)
putative"
/codon_start=1
/protein_id="BAC33256.1"
/db_xref="GI:26339170"
/translation="MRGRVTVANVTVKVPTLKRILPKLDVLEGENAVLIVETQAGV
OGCWSRDEGLPDTQSCSGHMRALVLPVTRDAGEITFSLGNSRTTLRVKCVKH
SPFGPVNMFVKGQKNKLVITWKPPEPPETSYFLRLEROEVGSDWIOCFSEIKAG
AVPEVDCVPTEGDFHFRICTVSEHGRSPHVFNFGSAHLVPTARLVSGLEDVQVYDGE
DAFSLDLSALIQGSFWFLNGEQNEPQVPCALRYIEQKGLHRLTLQAVKIR
DSGALVFCSPGVDSAAALITQESSVHILSPQDKVSLTFTTSERVLTCELSDVDFPA
TWYKDGQVSESESLIVKTEGRKRLILPEAQVRDSEFECECTEGVSFAFGVTVDPP
VHYNPOHVFVHALTSECVRILTCEVDREDTTVHYKDGQVSESDIIVLENKQPHRR
LVLPAAKSDGGEFCQVAGDERAFTVTITDVFVSNVTPSSEVHVAARLVRLVLTCE
LCRPWAEVRWTKDGEVVEVSPALLLEKEDTIRLVLPVQLEDSEYLCEIHDSASF
TITVTEPVRIIYPQDEVTLHVASLECVLTCELSREDAPVRYKDGLEVESEBALVL
QSDGPRRLVLPAAQPEDGEFVCDAGDSDSAFTVTVPPIVHPAARSLDLQFGA
FGHVELRCEVAPAGSVRWYKDGLEVEVSDALQIAGRPARTLTLPHQPEDAGEVYC
ETRDENVTVSLAELPVQFLAPEANPLCVRGEPPVLSCELSRASAQVFWSHNS
PVQOGEGLERAEGRPRILICQADLAHTGVITCOSGASPGAPSLSNVQVDELPPVK
LVSELTPLVHEGDATFQCEVSPDPAEVTWLRNGAVITAGPQLEVMVQNGSRLLIIR
GCQLKDGTVTARAGADTSARLHVRGSV"

CDS

14.1%; Score 140.8; DB 3; Length 4130;
Best Local Similarity 47.5%; Pred. No. 4.5e-22;
Matches 452; Conservative 0; Mismatches 497; Indels 3; Gaps 1;
Qy 1 AAAGATGGGAAGAAGCTCAGCTCCAGCTTGAAGAGTGCATGTAGAGGCCCAAGGCTGCAGA 60
Db 2217 AAAGATGGGAGGAGGTAGTGAGAGCCAGCAGCTGCTCCTGGAGAGGAGACCATC 2276
Qy 61 CGGAGGCTGGTGTGCACAGCAGCGAGGCAAGACGATGCCGGGGACTACAGCTCGAGGCC 120
Db 2277 CGCCGCTGGTGTGCTGCCCTCTGTCCAGCTTGAGGATCTTGGCGAGTACCTGTGTGAAATC 2336
Qy 121 AGGGCCACAGAGGGTCTCCTCTCCGCTGCACATCACAGAGCCCAAGATGATGTTTTCAAAG 180

ORIGIN

Query Match 14.1%; Score 140.8; DB 3; Length 4130;
Best Local Similarity 47.5%; Pred. No. 4.5e-22;
Matches 452; Conservative 0; Mismatches 497; Indels 3; Gaps 1;
Qy 1 AAAGATGGGAAGAAGCTCAGCTCCAGCTTGAAGAGTGCATGTAGAGGCCCAAGGCTGCAGA 60
Db 2217 AAAGATGGGAGGAGGTAGTGAGAGCCAGCAGCTGCTCCTGGAGAGGAGACCATC 2276
Qy 61 CGGAGGCTGGTGTGCACAGCAGCGAGGCAAGACGATGCCGGGGACTACAGCTCGAGGCC 120
Db 2277 CGCCGCTGGTGTGCTGCCCTCTGTCCAGCTTGAGGATCTTGGCGAGTACCTGTGTGAAATC 2336
Qy 121 AGGGCCACAGAGGGTCTCCTCTCCGCTGCACATCACAGAGCCCAAGATGATGTTTTCAAAG 180

Search completed: March 21, 2005, 14:13:33
Job time : 3669.47 secs

Db 2337 CATGATGAGTCGGCTTCCTTCCATCACCGTCACAGAGCCCTGTCGCGATCATATAC 2396
Qy 181 GAGCAGTCAGTGCATATAGGTGAGGCTGAGCGGGGGCCAGTCATCTCAGCTGT 240
Db 2397 CCCAGAGCAGAGGTGACCT---TACAGCCGTGAGTTTGAATGTGTGCTCCTCACCTGT 2453
Qy 241 GAGGTGGCCAGGCCAGAGCGAGGTGACGTGTGTACAGGATGGGAAGCAAGCTGAGCTCC 300
Db 2454 GAGCTGTCTAGAGAGGATGCTCCTGTGACGTGTGTACAGGATGGTTAGAGGTGAGGAG 2513
Qy 301 AGCTCAAAAGTGGCGCATGGAGGTCAAAGGGTGCACACGAGGCTGTGTGCTGCCACAGCGC 360
Db 2514 AGTGAAGCCCTGTGTGCTCCAGAGCGATGGGCTCTGTCGCGCTGTGTGCTGTGCTGCC 2573
Qy 361 GGCAAAACAGATGCTGGGGAGTACAGCTGTGAGGCTGGGGCCAGAGAGTCTCCTTCCAC 420
Db 2574 CAGCCAGAGACGGGGCGAGTTTGTGTGTGATCTGGGGATGATTGAGCCTTCTTCACT 2633
Qy 421 CTGCACATCACAGAGCCCAAGGGGTGTTCGCAAGGAGCAGTCAGTGCATATGAGGTG 480
Db 2634 GTCACTGTCAAGCTCCACAGAAAGGATTGTGCACCCAGCGGCCCGATCCCTGATTTG 2693
Qy 481 CAGCTGAGCGGGGGACCACTGCCATCTGAGCTGTGAGGTGGCCACGCCCAAGCGAG 540
Db 2694 CAGTTTGGGGCTCCAGGACACGTGGAGCTACGCTGCGAAGTGGCCCGGCTGGGTCTCAG 2753
Qy 541 GTGACGTGTACAGGACGAGGAGAGAGCTGAGCTCCAGCTCAAAAGTACCATGGAGGTC 600
Db 2754 GTGGCTGTACAAAGGATGGCTAGAGGTAGAGGTGTACATGGCGTGCAGCTGGGTGCT 2813
Qy 601 AAGGGCTGCACAGAAAGGCTGTGTAGTCAGCAGGTGGGCAAGCAGATGCTGGGGAGTAC 660
Db 2814 GAGGGGCTGCTCCGCACTCTCACCTGCCCCACGCCCTGAGGATGCCGGGAGTAT 2873
Qy 661 AGCTGCGAGCTGGGGCCAGAGAGTCTCCTTTCACTGCACATCACAGAGCCCAAGGCA 720
Db 2874 GTATGTGAGACCCGAGATGAGGCTGTCACTTCAACGTCAGCCTGGGCTGAGCTTCCGGTG 2933
Qy 721 GTGTTTCCCAAGGAGCAGTTTGGTGCATAATGAGTGGGACTGAGGCGAGGGCCAGTGCC 780
Db 2934 CAGTTTCTGGCTCCAGAGGCGAGCCCCAATCCGCTCTGCGTGTCTTGGGGAGCCCTG 2993
Qy 781 ACATGAGCTGTGAGGTGGCCAGAGGCCAGACAGAGTGCAGTGTGTTTCAAGAGTGGAAAG 840
Db 2994 GTGTGAGCTGTGAGCTGTCCGAGCAAGCGCACAGTGTCTTGGAGCCACAATGGGAGC 3053
Qy 841 AAGCTGAGCTCCAGTTTCAAGTCCGATAGAGGCTGGGGCTGCATGCGGCGAGCTGGTG 900
Db 3054 CCGGTGAGCAGGGTGAAGACTAGAGCTCGAGCTGAGGGTCTCTCCAGAAATCCTCTGC 3113
Qy 901 GTGCAGCAGGCGAGGCCAGGAGATGCTGGGGAGTACACCTGTGAGGCTGGGG 952
Db 3114 ATCCAGCAGCAGACCTCGCTCACAGAGTGTCTACACTTGCCTTGCCTTCCAAATCTGGGG 3165

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:15:10 ; Search time 581.044 Seconds
(without alignments)
10198.307 Million cell updates/sec

Title: US-10-077-130-4_COPY_22500_23500

Perfect score: 1001

Sequence: 1 ggccagtcagtgacactggc.....ctgtcaagtcgtggacct 1001

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq_16Dec04.*

1: geneseqn1980s.*

2: geneseqn1990s.*

3: geneseqn2000s.*

4: geneseqn2001as.*

5: geneseqn2001bs.*

6: geneseqn2002as.*

7: geneseqn2002bs.*

8: geneseqn2003as.*

9: geneseqn2003bs.*

10: geneseqn2003cs.*

11: geneseqn2003ds.*

12: geneseqn2004as.*

13: geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	3225	5	AAD26467 Human kin
2	1001	100.0	4175	6	AAC62285 cDNA enco
3	1001	100.0	4936	4	AHA46904 cDNA enco
4	1001	100.0	5454	6	ABD38865 Human kin
5	1001	100.0	8106	8	ABX11641 Human ser
6	1001	100.0	24120	8	ABX11642 Human ser
7	999.4	99.8	5007	5	AAC62286 cDNA enco
8	999.4	99.8	5207	6	AAL43908 Human kin
9	999.4	99.8	5207	6	AAL43908 Human kin
10	999.4	99.8	7928	5	AAC62287 cDNA enco
11	385	38.5	876	7	ADR41256 Human CD-
12	164.8	16.5	182	3	AAC25654 Human sec
13	151.6	15.1	9930	10	ADE47671 Human NOV
14	151.6	15.1	9930	12	ADJ78941 Human NOV
15	150	15.0	7710	4	AAF44662 Novel pro
16	150	15.0	7710	12	ADI29360 Human MAR
17	148.4	14.8	7789	6	AD30565 Human kin
18	148.4	14.8	10662	10	ADC99127 Human KPP
19	146.8	14.7	9698	10	ADE47675 Human NOV
20	146.8	14.7	9698	12	ADJ78945 Human NOV

21	146.8	14.7	10122	10	ADB47673 Human NOV
22	146.8	14.7	10122	12	ADJ78943 Human NOV
23	140.4	14.0	9807	10	ADB79958 Human kin
24	105.6	10.5	62805	10	ADB79960 Human kin
25	97.8	9.8	860	10	ADB47677 Human NOV
26	97.8	9.8	860	12	ADJ78947 Human NOV
27	92	9.2	1365	12	ADO71726 Human gen
28	92	9.2	2055	13	ADN60271 Human ZIP
29	92	9.2	2055	13	ADR83403 Human ZIP
30	92	9.2	2079	4	AAK94258 Human ful
31	92	9.2	2079	12	ADL30841 Full leng
32	92	9.2	2079	13	ADQ68012 Recombina
33	92	9.2	2104	12	ADQ84258 Human tum
34	92	9.2	2104	13	ADQ85644 Human tum
35	92	9.2	2104	13	ADQ86740 Human tum
36	92	9.2	2224	4	AAH16158 Human cDN
37	92	9.2	2224	5	AAH78068 Nucleotid
38	92	9.2	2226	10	ADD29770 Human tum
39	90.4	9.0	757	4	AAK93262 Human cDN
40	90.4	9.0	757	4	AAK91856 Human cDN
41	90.4	9.0	757	12	ADL29689 5' end of
42	90.4	9.0	757	12	ADL28283 5' end of
43	90.4	9.0	2105	6	ABV72291 Nucleotid
44	90.4	9.0	2105	13	ADN60267 Human dea
45	90.4	9.0	2105	13	ADN60269 Human ZIP

ALIGNMENTS

RESULT 1

AAD26467

ID AAD26467 standard; cDNA; 3225 BP.

XX AC AAD26467;

DT 26-MAR-2002 (first entry)

XX DE Human kinase PKIN-20 cDNA.

Human; kinase; PKIN-20; cancer; leukaemia; adenocarcinoma; osteoporosis; immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease; Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia; allergy; asthma; adult respiratory distress syndrome; multiple sclerosis; autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis; Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis; rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome; hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris; cardiovascular disease; hypertension; vasculitis; myocarditis; obesity; congestive heart failure; ischaemic heart disease; lung tumour; gout; fatty liver; Niemann-Pick's disease; gene therapy; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT CDS 526..3141

FT /*tag= a

FT /product= "Human PKIN-20 protein"

XX WO200196547-A2.

XX PD 20-DEC-2001.

XX 14-JUN-2001; 2001WO-US019444.

XX 15-JUN-2000; 2000US-0212073P.

XX 23-JUN-2000; 2000US-0213467P.

XX 30-JUN-2000; 2000US-0215651P.

XX 07-JUL-2000; 2000US-0216605P.

XX 13-JUL-2000; 2000US-0218372P.

XX 25-AUG-2000; 2000US-0228056P.

XX (INCY-) INCYTE GENOMICS INC.

XX	
PI	Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
PI	Gandhi AR, Tribouley CM, Walia NK, Yao MG, Lu DAM, Greenwald SR;
PI	Ranunkar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;
PI	Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
PI	Lo TP, Khan F, Recipon SA, Azinzai Y, Policky JL, Ding L;
PI	Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
XX	
DR	WPI; 2002-090207/12.
DR	P-PSDB; AA616274.
DR	

PT New polypeptides, useful for diagnosing, treating or preventing disorders of growth and development, cardiovascular and lipid, and diseases such as cancer, comprise human kinase polypeptides.

Claim 5; Page 191; 197pp; English.

The invention relates to human kinase PKIN proteins and their corresponding cDNAs. A composition containing PKIN agonist is useful for treating a disease or condition associated with decreased expression of PKIN and a composition comprising PKIN antagonist is useful for treating a disease or condition associated with overexpression of PKIN. The disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, Hodgkin's disease), immune disorder (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease, atherosclerosis, allergies, adult respiratory distress syndrome, autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease, osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis, bacterial, parasitic, fungal, viral, protozoal and helminthic infections) growth and development disorders (arteriosclerosis, cirrhosis, hepatitis, Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts), cardio vascular disease (arteriovenous fistula, hypertension, vasculitis, aneurysms, congestive heart failure, angina pectoris, myocarditis, ischaemic heart disease, chronic bronchitis, lung tumours), lipid disorder (fatty liver, Fabry's disease, Niemann-Pick's disease, hypcholesterolaemia, obesity). PKIN DNA is useful for assessing toxicity of a test compound and in gene therapy. The present sequence is human PKIN-20 cDNA

Sequence 3225 BP; 596 A; 1104 C; 1017 G; 508 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 6; Length 3225;

Best Local Similarity 100.0%; Pred. No. 1.9e-197;

QY 1 GGCCAGTCAGTGACACTGGCCTGCCAGGTGTGAGCCAGCCAGCTGCCACGCCCACCTGG 60

Db 1663 GGCCAGTCAGTGACACTGGCCCTGCCAGGTGTAGCCAGCCAGCTGCCAGGCCACCTGG 1722

QY 61 AGCAAAGACGGAGCCCCCTGGAGAGCAGCCGTGCTCTCATCTCTGCCACCCCTCAAG 120

Db 1723 AGCAAAGACGGAGCCCCCTGGAGAGCAGCCGTCTCATCTCTGCCACCCCTCAAG 1782

[illegible]

Db
1783 AACTTCCAGCTTCTGACCATCCTGGTGTTGGTGCTGAGGACCTGGGTGTACACCTGC 1842

181 AGCGTGAGCAATGCGCTGGGACAGTGACCAACGGCGTCTCCGGAAGCAGAGCGC 240

DB 1843 AGCGTGAGCAATGCGCTGGGGACAGTGACCACCAACGGGGCTCTCCGGAAGGCAGAGCGC 1902

QY 241 CCCCATCGCCGATATCGGGAGGTGTACGCGGATGGGGTGTCTGGTC 300

525 1903 CCCCATGCCCCGGATAACGGGAGGCTGACGCGGATGGGGTGCCTGGTC 1962

5' TGGTTCCTGTTCCATACCGCCCTGTGACCTACATTGTGCAGTGCAAGCCTAGTAGGC 360

2022

Db	2023	GGCAGCTGGACCACACTGGGCTCCGACATCTTTGACTGTGCTACCTGACCAAGCTC	2082
Qy	421	TCCGGGGTGGCACCTTACACTTCGCGACGGCATGTGTCAGCAAGCAGGATGGGTCCC	480
Db	2083	TCCGGGGTGGCACCTTACACTTCGCGACGGCATGTGTCAGCAAGCAGGATGGGTCCC	2142
Qy	481	TACAGCAGCCCTCGGAGCAAGTCTCTCTGGGAGGCCACGACCTGGCTCTGAGGAG	540
Db	2143	TACAGCAGCCCTCGGAGCAAGTCTCTCTGGGAGGCCACGACCTGGCTCTGAGGAG	2202
Qy	541	GAGGCCAGGGGGGTGACGCCAACCCCTGCCACAGCAAAAGCTTCGATTCAGACA	600
Db	2203	GAGGCCAGGGGGGTGAGCCAAACCCCTGCCACAGCAAAAGCTTCGATTCAGACA	2262
Qy	601	CAGATCCAGAGGGGCGCTTCACGCTGGTGGCGCAATGCTGGAGAAAGCCAGCGGGCGG	660
Db	2263	CAGATCCAGAGGGGCGCTTCACGCTGGTGGCGCAATGCTGGAGAAAGCCAGCGGGCGG	2322
Qy	661	GCCTGGCGCCGAAGATCATCTCCCTACACCCCAAGCAAGACAGCTGTGGCGGAA	720
Db	2323	GCCTGGCGCCGAAGATCATCTCCCTACACCCCAAGCAAGACAGCTGTGGCGGAA	2382
Qy	721	TACGAGGCCCTCAAGGGCTGGGCCACCGCACCTGGCCACGTCGACGACGCTACCTC	780
Db	2383	TACGAGGCCCTCAAGGGCTGGGCCACCGCACCTGGCCACGTCGACGACGCTACCTC	2442
Qy	781	AGCCCCGGCACTGTGTCTCATCTTGGAGCTGTGCTCTGGGCCCGAGCTGTCCCTGCG	840
Db	2443	AGCCCCGGCACTGTGTGTCTCATCTTGGAGCTGTGCTCTGGGCCCGAGCTGTCCCTGCG	2502
Qy	841	CTGGCCGAGGGGCTTCCTACCTCAGATCCGAGGTGAAGGACTACTGTGGCAGATGTTG	900
Db	2503	CTGGCCGAGGGGCTTCCTACCTCAGATCCGAGGTGAAGGACTACTGTGGCAGATGTTG	2562
Qy	901	AGTGCCACCCAGTACTCTGCACAAACAGCACATCTCTGCACCTGGACCTGAGGTCGAGAAC	960
Db	2563	AGTGCCACCCAGTACTCTGCACAAACAGCACATCTCTGCACCTGGACCTGAGGTCGAGAAC	2622
Qy	961	ATGATCATCAGCAATACAACTGCTCAAGGTCGTGGACCT	1001
Db	2623	ATGATCATCAGCAATACAACTGCTCAAGGTCGTGGACCT	2663

RESULT 2

AAC62285
 ID AAC62285 standard; cDNA; 4175 BP.

AAC AAC62285;

19-MAR-2001 (first entry)

DE cDNA encoding a human signal transduction polypeptide.

Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
 KW congestive heart failure; dilated congestive cardiomyopathy;
 KW hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
 KW mitral valve disease; aortic valve disease; tricuspid valve disease;
 KW myocardial infarction; cardiac arrhythmia; arteriosclerosis;
 KW atherosclerosis; cardiac tumour; microbial infection; ss.

OS Homo sapiens.

XX	Key	Location/Qualifiers
FH		

```

FT      CDS
I: .4036
/*tag= a

```

XX
F1
/product= "signal transduction polypeptide H19G5XX
XX-TOCCOAOZOU

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841

XX
XX

PI Plowman G, Whyte D, Manning G, Sudarsanam S, Martinez R;
 DR WPI: 2001-476202/51.
 DR P-PSDB; AAB85504.
 XX Kinase polypeptides useful for treating cancers, Alzheimer's disease,
 PT viral infections, diabetes, obesity, organ transplant rejection and
 PT rheumatoid arthritis.
 XX Example 1; Page 204-205; 218pp; English.
 XX The invention provides human protein kinases and protein kinase-like
 CC enzymes and polynucleotides encoding the polypeptides. The kinase
 CC polypeptides and their modulators are useful for treating a disease or
 CC disorder such as cancer, immune-related diseases, cardiovascular disease,
 CC brain or neuronal-associated disease and metabolic disorders, including
 CC cancers of tissues, cancers of hematopoietic origin, diseases of the
 CC central nervous system, diseases of the peripheral nervous system,
 CC Alzheimer's disease, Parkinson's disease, multiple sclerosis, amyotrophic
 CC lateral sclerosis, viral infections, infections caused by prions,
 CC bacteria and fungi, ocular diseases, migraines, pain, sexual dysfunction,
 CC mood disorders, attention disorders, cognition disorders, hypotension,
 CC hypertension, psychotic disorders, neurological disorders, dyskinesias,
 CC metabolic disorders, and organ transplant rejection. They are also useful
 CC for treating rhinitis, autoimmunity, atherosclerosis, psoriasis,
 CC osteoarthritis, asthma, chronic inflammatory pelvic disease, chronic
 CC inflammatory bowel disease, rheumatoid arthritis, metabolic disorders
 CC such as diabetes, obesity, cardiovascular diseases such as reperfusion
 CC injury, coronary thrombosis, clotting disorders and atherosclerosis,
 CC ocular diseases such as glaucoma, retinopathy and macular degeneration,
 CC psychiatric and neurological disorders such as anxiety, schizophrenia,
 CC dementia, manic depression, etc. The polynucleotides are useful in gene
 CC therapy techniques to treat the above mentioned disorders. Sequences
 CC AAH46891-46922 represent human protein kinases encoding cDNA molecules
 XX
 SQ Sequence 4936 BP; 923 A; 1695 C; 1524 G; 794 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 4; Length 4936;
 Best Local Similarity 100.0%; Pred. No. 2e-197;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GGCAGTCAGTGACACTGGCTGCGAGTGTGAGGAGCAGCGGTGCTCATCTCTGCCACCTCAAG 60
 DB 3379 GGCACGTGAGTGACACTGGCTGCGAGTGTGAGGAGCAGCGGTGCTCATCTCTGCCACCTCAAG 3438
 QY 61 AGCAAGACGGAGCCCTGCGAGAGCAGCGGTGCTCATCTCTGCCACCTCAAG 120
 DB 3439 AGCAAGACGGAGCCCTGCGAGAGCAGCGGTGCTCATCTCTGCCACCTCAAG 3498
 QY 121 AACTTCCAGCTTGTGACCATCTGGTGTGTTGGTGTGAGGACCTGGGTGTGTACACCTGC 180
 DB 3499 AACTTCCAGCTTGTGACCATCTGGTGTGTTGGTGTGAGGACCTGGGTGTGTACACCTGC 3558
 QY 181 AGCGTGACATCGCTGGGACAGTGAACACACGCGGCTCTCCGGAAGGAGAGCGC 240
 DB 3559 AGCGTGACATCGCTGGGACAGTGAACACACGCGGCTCTCCGGAAGGAGAGAGCGC 3618
 QY 241 CCCTCATCTTCCGATCCCGGATATCGGAGAGTGTACGCGGATGGGGTCTGCTGTC 300
 DB 3619 CCCTCATCTTCCGATCCCGGATATCGGAGAGTGTACGCGGATGGGGTCTGCTGTC 3678
 QY 301 TGGAGCCCGTGGAACTCTACGGCCCTGTGACCTACATTTGCGAGTGCAGCTAGAAGGC 360
 DB 3679 TGGAGCCCGTGGAACTCTACGGCCCTGTGACCTACATTTGCGAGTGCAGCTAGAAGGC 3738
 QY 361 GGCAGCTGGAACCACTGGCTCGGACATCTTGTGCTGCTCTACCTGACGAGCAAGTC 420
 DB 3739 GGCAGCTGGAACCACTGGCTCGGACATCTTGTGCTGCTCTACCTGACGAGCAAGTC 3798
 QY 421 TCCGGGGTGGCACTTACACCTTCCGACCGGCTGTGTGACGAGGAGGAGTGGGTCCC 480
 DB 3799 TCCGGGGTGGCACTTACACCTTCCGACCGGCTGTGTGACGAGGAGGAGTGGGTCCC 3858

QY 481 TACAGCAGCCCTCGAGCAAGTCTCTCTGGAGGGGCCAGCCACCTCTGGCCTCTGAGGAG 540
 DB 3959 TACAGCAGCCCTCGAGCAAGTCTCTCTGGAGGGGCCAGCCACCTCTGGCCTCTGAGGAG 3918
 QY 541 GAGAGCAGGGGGCGGTGAGCCCAACCCCTGCTGCCAGCAAAAGACTTCTGCANTTCCAGACA 600
 DB 3919 GAGAGCAGGGGGCGGTGAGCCCAACCCCTGCTGCCAGCAAAAGACTTCTGCANTTCCAGACA 3978
 QY 601 CAGATCCAGAGGGGGCGGTGAGCCCAACCCCTGCTGCCAGCAAAAGACTTCTGGGAGAGGCCGCGGG 660
 DB 3979 CAGATCCAGAGGGGGCGGTGAGCCCAACCCCTGCTGCCAGCAAAAGACTTGGGGAGAGGCCGCGGG 4038
 QY 661 GCGTGGCGCGCAAGATCATCCCTTACACCCCAAGGACAGACAGCAGTGTCTGCGGAA 720
 DB 4039 GCGTGGCGCGCAAGATCATCCCTTACACCCCAAGGACAGACAGCAGTGTCTGCGGAA 4098
 QY 721 TACGAGCCCTCAAGGGCGCTGCGCCACCCGACCTGCGCCCAAGTGTGACGAGCCTACTCTC 780
 DB 4099 TACGAGCCCTCAAGGGCGCTGCGCCACCCGACCTGCGCCCAAGTGTGACGAGCCTACTCTC 4158
 QY 781 AGCCCCCGGACCTGCTGCTCATCTTGGAGCTGTCTTGGGCCGAGCTGCTCCCTGC 840
 DB 4159 AGCCCCCGGACCTGCTGCTCATCTTGGAGCTGTCTTGGGCCGAGCTGCTCCCTGC 4218
 QY 841 CTGGCCGAGAGGGCGCTCTTACTCAGAAATCCGAGTGAAGACTACCTGTGCGAGATGTTG 900
 DB 4219 CTGGCCGAGAGGGCGCTCTTACTCAGAAATCCGAGTGAAGACTACCTGTGCGAGATGTTG 4278
 QY 901 AGTGCCACCCAGTACTGTGCAACACAGCACAATCTGACCTGACCTGAGGTGCGAGAAC 960
 DB 4279 AGTGCCACCCAGTACTGTGCAACACAGCACAATCTGACCTGAGGTGCGAGAAC 4338
 QY 961 ATGATCATACCGAATACACCTGCTCAAGGTGCGACCT 1001
 DB 4339 ATGATCATACCGAATACACCTGCTCAAGGTGCGACCT 4379

RESULT 4
 AAD38865
 ID AAD38865 standard; cDNA; 5454 BP.
 AC AAD38865;
 XX
 DT 23-SEP-2002 (first entry)
 XX
 DE Human kinase (PKIN)-22 cDNA.
 XX
 KW Human; kinase; PKIN; cancer; immune system disorder; atherosclerosis;
 KW acquired immune deficiency syndrome; AIDS; Addison's disease; allergy;
 KW asthma; multiple sclerosis; psoriasis; arteriosclerosis; cirrhosis;
 KW development; hepatitis; cardiovascular; hypertension; drug screening;
 KW myocardial infarction; Goodpasture's syndrome; lipid disorder; growth;
 KW fatty liver; Gaucher's disease; Niemann-Pick's disease; anorectic;
 KW hypercholesterolemia; obesity; gene therapy; cytostatic; anti-hiv;
 KW neuroprotective; hepatotropic; hypotensive; cardiant; nephrotropic;
 KW hyperlipidaemia; enzyme; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 CDS 373..5370
 FT /*tag= a
 FT /product= "Human kinase (PKIN)-22"
 XX
 PN WO200233099-A2.
 XX
 PD 25-APR-2002.
 XX
 PF 20-OCT-2001; 2001WO-US047728.
 XX
 PR 20-OCT-2000; 2000US-0242410P.
 PR 27-OCT-2000; 2000US-0244068P.
 PR 03-NOV-2000; 2000US-0245708P.

PR 09-NOV-2000; 2000US-0247672P.
 PR 16-NOV-2000; 2000US-0249565P.
 PR 22-NOV-2000; 2000US-0252730P.
 PR 01-DEC-2000; 2000US-0250807P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Gururajan R, Baughn MR, Wallia NK, Elliott VS, Xu Y, Arvizu C;
 PI Yao MG, Ramkumar J, Ding L, Tang Y, Hafalia AJA, Nguyen DB;
 PI Gandhi AR, Lu Y, Yue H, Burford N, Bandman O, Tribouley CM, Lal PG;
 PI Recipon SA, Lu DAM, Borowsky ML, Thornton M, Swarner A;
 PI Thangavelu K, Khan FA, Ison CH;
 XX
 XX WPI; 2002-454603/48.
 DR P-PSDB; AAE24151.
 DR
 XX
 XX New human kinase polypeptide, for diagnosing, preventing and treating
 PT cancer, immune system disorders, growth and development disorders,
 PT cardiovascular disorders and lipid disorders.
 PT
 XX Claim 5; Page 209-210; 210pp; English.
 PS
 XX The invention relates human kinases (PKIN) and their corresponding
 CC nucleic acid sequences. PKIN and its DNA are useful for diagnosing,
 CC treating and preventing cancer, an immune system disorder (e.g., acquired
 CC immune deficiency syndrome (AIDS), Addison's disease, allergy, asthma,
 CC atherosclerosis, multiple sclerosis, psoriasis), disorders affecting
 CC growth and development (e.g., arteriosclerosis, cirrhosis, hepatitis),
 CC cardiovascular disorder (e.g., hypertension, myocardial infarction,
 CC Goodpasture's syndrome), and a lipid disorder (e.g., fatty liver,
 CC Gaucher's disease, Niemann-Pick's disease, hypercholesterolemia,
 CC hyperlipidaemia, obesity), and for assessing the effects of exogenous
 CC compounds. Anti-PKIN antibody is useful in a diagnostic test for a
 CC condition or a disease associated with the expression of PKIN in a
 CC biological sample. A composition comprising PKIN or an agonist or
 CC antagonist of PKIN is useful for treating a disease or condition
 CC associated with decreased or increased expression of functional PKIN.
 CC PKIN is useful in a number of drug screening techniques and to analyse
 CC the proteome of a tissue or cell type. PKIN DNA is useful for creating
 CC knockin humanised animals or transgenic animals to model human diseases,
 CC and in somatic or germline gene therapy. The present sequence is human
 CC PKIN CDNA
 XX
 XX Sequence 5454 BP; 1019 A; 1853 C; 1686 G; 896 T; 0 U; 0 Other;
 SQ

Query Match 100.0%; Score 1001; DB 6; Length 5454;
 Best Local Similarity 100.0%; Pred. No. 2e-197;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGCAGTCAGTGACACTGGCTGCGAGGTGTACGCCAGCCAGCGTGGCCAGCCACCTGG 60
 Db 3892 GGCAGTCAGTGACACTGGCTGCGAGGTGTACGCCAGCCAGCGTGGCCAGCCACCTGG 3951

Qy 61 AGCAAGACGAGCGCCCTCGAGAGCAGCAGCGTGTCTCATCTCTGCCACCTCAAG 120
 Db 3952 AGCAAGACGAGCGCCCTCGAGAGCAGCAGCGTGTCTCATCTCTGCCACCTCAAG 4011

Qy 121 AACTTCCAGCTTCTGACCATCTCTGGTGTGGCTGAGGACCTGGGTGTGTACACCTGC 180
 Db 4012 AACTTCCAGCTTCTGACCATCTCTGGTGTGGCTGAGGACCTGGGTGTGTACACCTGC 4071

Qy 181 AGCGTGAGCAATGGCTGGGACAGTGACACACAGCGGCGTCTCCGGAAGCGCAGAGCGC 240
 Db 4072 AGCGTGAGCAATGGCTGGGACAGTGACACACAGCGGCGTCTCCGGAAGCGCAGAGCGC 4131

Qy 241 CCCTCATCTTCGCAATCGCGGATATCGGGAGGTGTACGGGATGGGGTCTCTGGTC 300
 Db 4132 CCCTCATCTTCGCAATCGCGGATATCGGGAGGTGTACGGGATGGGGTCTCTGGTC 4191

Qy 301 TGGAGCCCGTGGGAATCTACGGCCCTGTGACCTACATTGTGCAGTGCAGCTTAGAAGGC 360
 Db 4192 TGGAGCCCGTGGGAATCTACGGCCCTGTGACCTACATTGTGCAGTGCAGCTTAGAAGGC 4251

Qy 361 GGCAGCTGGACACACACTGGGCTCCGACATCTTTGACTGCTGCTACCTGACCAAGCTC 420
 Db 4252 GGCAGCTGGACACACACTGGGCTCCGACATCTTTGACTGCTGCTACCTGACCAAGCTC 4311

Qy 421 TCCCGGGGTGGCACTTACACCTTCGCAAGCGCATGTGTACGAAAGGCAAGGATGGGTCCC 480
 Db 4312 TCCCGGGGTGGCACTTACACCTTCGCAAGCGCATGTGTACGAAAGGCAAGGATGGGTCCC 4371

Qy 481 TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGGCCAGCCACCTGGCTCTTGAGAG 540
 Db 4372 TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGGCCAGCCACCTGGCTCTTGAGAG 4431

Qy 541 GAGAGCCAGGGCGGTGAGCCCAACCTGCGCCAGCACAAGAGCTTCCGATTCAGACA 600
 Db 4432 GAGAGCCAGGGCGGTGAGCCCAACCTGCGCCAGCACAAGAGCTTCCGATTCAGACA 4491

Qy 601 CAGATCCAGAGGGGCCGCTTCAGGCGTGGTGGGCAATGCTGGGAGAAAGCCAGCGGGCGG 660
 Db 4492 CAGATCCAGAGGGGCCGCTTCAGGCGTGGTGGGCAATGCTGGGAGAAAGCCAGCGGGCGG 4551

Qy 661 GCGTGGCGCCCAAGATCATCTCCCTTACCAACCCCAAGGACAAAGACGAGTCTGCGGAA 720
 Db 4552 GCGTGGCGCCCAAGATCATCTCCCTTACCAACCCCAAGGACAAAGACGAGTCTGCGGAA 4611

Qy 721 TACAGAGCCCTCAGGGGCTGGCCACCCGACCTGGCCAGCTGCGCCAGCTGACGAGCTTACCTC 780
 Db 4612 TACAGAGCCCTCAGGGGCTGGCCACCCGACCTGGCCAGCTGCGCCAGCTGACGAGCTTACCTC 4671

Qy 781 AGCCCCGGCACTGGTGTCTCATCTGGAGTGTGCTCTGGGCGCGAGCTGTCCCTTCG 840
 Db 4672 AGCCCCGGCACTGGTGTCTCATCTGGAGTGTGCTCTGGGCGCGAGCTGTCCCTTCG 4731

Qy 841 CTGCGCAGAGGGCTCTTACTCAGATCCGAGTGAAGACTACCTGTGCGAGATGTTG 900
 Db 4732 CTGCGCAGAGGGCTCTTACTCAGATCCGAGTGAAGACTACCTGTGCGAGATGTTG 4791

Qy 901 AGTGCACCCAGTACCTGCGACACACACACCTCTGCACCTGGACCTGAGTCCGAGAAC 960
 Db 4792 AGTGCACCCAGTACCTGCGACACACACACCTCTGCACCTGGACCTGAGTCCGAGAAC 4851

Qy 961 ATGATCATACCGAATACAACTCTCAAGGTCTGGACCT 1001
 Db 4852 ATGATCATACCGAATACAACTCTCAAGGTCTGGACCT 4892

RESULT 5
 ABX11641
 ID ABX11641 standard; cDNA; 8106 BP.
 XX
 AC ABX11641;
 XX
 DT 09-MAY-2003 (first entry)
 XX
 DE Human serine/threonine or protein kinase 59079, cDNA.
 XX
 KW Human; ss; gene; serine/threonine kinase; protein kinase; 59079;
 KW cardiovascular disease; heart failure; myocardial infarction;
 KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma;
 KW blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
 KW haemolytic anaemia; cellular proliferative disorder; cancer;
 KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
 KW psoriasis; inflammatory bowel disease; rheumatoid arthritis;
 KW multiple sclerosis.
 XX
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 5'UTR 1..71
 FT /*tag= a
 FT CDS 72..7964
 FT /*tag= b
 FT /product= "Kinase 59079"
 FT /note= "This CDS is specifically claimed in claim 2"

FT 3' UTR 7965.. .8106
 FT /*tag= c
 XX
 PN US2002168742-A1.
 XX
 PD 14-NOV-2002.
 XX
 PF 15-FEB-2002; 2002US-00077130.
 XX
 PR 15-FEB-2001; 2001US-0269201P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Kapeller-Iibermann R, Acton SL;
 XX
 DR WPI: 2003-298729/29.
 DR P-PSDB; ABG76186.
 XX
 PT Novel isolated human protein kinase, designated 59079 or 12599
 PT polypeptide, useful as diagnostic and therapeutic agents for preventing
 PT cardiovascular diseases, proliferative disorders, and protein kinase
 PT disorders.
 XX
 PS Claim 2; Page 39-48; 119pp; English.
 XX
 CC The invention relates to an isolated human serine/threonine or protein
 CC kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule
 CC comprising at least 85% identity to the nucleic acids appearing as
 CC ABX11641 and ABX11642 or their complement, a naturally occurring variant
 CC of the kinases or their fragments. Also included are a non-human host
 CC cell containing the nucleic acids, an antibody specific for the proteins,
 CC identifying a compound which binds to the kinase (by contacting the
 CC kinase or a cell expressing the kinase with a test compound and
 CC determining whether the kinase binds to the test compound) and modulating
 CC the activity of kinase using the identified compound. The kinases and
 CC their encoding nucleic acids are useful as diagnostic and therapeutic
 CC agents for preventing a disease or condition associated with an aberrant
 CC or unwanted 59079 or 12599 activity in a subject, including
 CC cardiovascular diseases such as heart failure, and myocardial infarction;
 CC disorders involving blood vessels such as atherosclerosis, and Kaposi's
 CC sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia,
 CC Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders
 CC such as cancer; and protein kinase disorders such as autoimmune
 CC disorders, diabetes mellitus, psoriasis, inflammatory bowel disease,
 CC rheumatoid arthritis, and multiple sclerosis (many examples of diseases
 CC and disorders are included in the specification). The kinases, their
 CC encoding nucleic acids and antibodies are useful in screening assays,
 CC detection assays (e.g. forensic biology), and predictive medicine (e.g.
 CC diagnostic assays, prognostic assays, and monitoring clinical trials and
 CC pharmacogenomics). The kinases and their encoding nucleic acids are
 CC useful as query sequences to perform a search against public databases to
 CC identify other family members or related sequences. The present sequence
 CC encodes the kinase 59079
 XX
 SQ Sequence 8106 BP; 1603 A; 2684 C; 2530 G; 1289 T; 0 U; 0 Other;

Db	6666	AGCGTGAGCAATCGCTGGGGACAGTGACCAACACGCGCGTCTCCCGGAAGGACAGAGCGC	6725
Qy	241	CCCTCATCTTCGCCCATGCGCGGATATCGGGAGGTATGACGGGATGGGTGCTGCTGGTC	300
Db	6726	CCCTCATCTTCGCCCATGCGCGGATATCGGGAGGTATGACGGGATGGGTGCTGCTGGTC	6785
Qy	301	TGGAAGCCCGTGGAAATCCTACGGCCCTGTGACCTACATTGTGCAGTGCAGCCTAGAGGC	360
Db	6786	TGGAAGCCCGTGGAAATCCTACGGCCCTGTGACCTACATTGTGCAGTGCAGCCTAGAGGC	6845
Qy	361	GGCAGCTGGACCAACACTGGCCCTCCGACATCTTTTGACTGCTGCTACCTGACAGCAAGTC	420
Db	6846	GGCAGCTGGACCAACACTGGCCCTCCGACATCTTTTGACTGCTGCTACCTGACAGCAAGTC	6905
Qy	421	TCCCGGGGTGGCACTACACTTCGCGACGGCATGTGTACGAAAGGAGGAGGATGGGTCCC	480
Db	6906	TCCCGGGGTGGCACTACACTTCGCGACGGCATGTGTACGAAAGGAGGAGGATGGGTCCC	6965
Qy	481	TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCACCTGGGCTCTGAGGAG	540
Db	6966	TACAGCAGCCCTCGGAGCAAGTCTCTCTGGAGGGCCAGCCACCTGGGCTCTGAGGAG	7025
Qy	541	GAGAGCCAGGGGGGTGAGCCCAACCCCTGCCAGCAACAAAGACCTTCGCATTCACAGACA	600
Db	7026	GAGAGCCAGGGGGGTGAGCCCAACCCCTGCCAGCAACAAAGACCTTCGCATTCACAGACA	7085
Qy	601	CAGATCCAGAGGGGGCCCTTCAGCGTGTGGGCAATGCTGGGAGAGGCGGCGGGGG	660
Db	7086	CAGATCCAGAGGGGGCCCTTCAGCGTGTGGGCAATGCTGGGAGAGGCGGCGGGGG	7145
Qy	661	CGCTGCGCCGCAAGATCATCCCTTACACCCCAAGGACAAAGACAGCAGTGCTGCGCGAA	720
Db	7146	CGCTGCGCCGCAAGATCATCCCTTACACCCCAAGGACAAAGACAGCAGTGCTGCGCGAA	7205
Qy	721	TACAGAGCCCTCAAGGCGCTTGCGCCACCCGACCTGGCCAGCTGACGAGCCTACCTC	780
Db	7206	TACAGAGCCCTCAAGGCGCTTGCGCCACCCGACCTGGCCAGCTGACGAGCCTACCTC	7265
Qy	781	AGCCCCCGGCACTGGTGCTCATCTTGGAGCTGTGCTCTGGGCCGAGCTGCTCCCTGC	840
Db	7266	AGCCCCCGGCACTGGTGCTCATCTTGGAGCTGTGCTCTGGGCCGAGCTGCTCCCTGC	7325
Qy	841	CTGGCCGAGAGGGCTCTCTACTCAGAAATCCGAGTGAAGGACTACCTGTGGCAGATTTG	900
Db	7326	CTGGCCGAGAGGGCTCTCTACTCAGAAATCCGAGTGAAGGACTACCTGTGGCAGATTTG	7385
Qy	901	AGTGCCACCCAGTACTGTGCAACACAGCACAATCTGCACTGACCTGAGCTGCGAAGC	960
Db	7386	AGTGCCACCCAGTACTGTGCAACACAGCACAATCTGCACTGACCTGAGCTGCGAAGC	7445
Qy	961	ATGATCATCACCAGAAATCAACCTGCTCAAGGTGTGGACCT 1001	
Db	7446	ATGATCATCACCAGAAATCAACCTGCTCAAGGTGTGGACCT 7486	

RESULT 6
 ABX11642
 ID ABX11642 standard; cDNA; 24120 BP.
 XX
 AC ABX11642;
 XX
 DT 09-MAY-2003 (first entry)
 XX
 DE Human serine/threonine or protein kinase 12599, cDNA.
 XX
 KW Human; ss; gene; serine/threonine kinase; protein kinase; 12599;
 KW cardiovascular disease; heart failure; myocardial infarction;
 KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma;
 KW blood platelet disorder; thrombocytopaenia; leukaemia;
 KW haemolytic anaemia; cellular proliferative disorder; cancer;
 KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
 KW psoriasis; inflammatory bowel disease; rheumatoid arthritis;

OS	Homo sapiens.	
XX		
Key	Location/Qualifiers	
CDS	109..5106	
FT	/*tag= a	
FT	/product= "Human kinase protein"	
FT	/replace(311, C)	
FT	/*tag= c	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	/replace(311, T)	
FT	/*tag= b	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	/replace(1741, T)	
FT	/*tag= d	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	/replace(2714, C)	
FT	/*tag= e	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	/replace(2745, T)	
FT	/*tag= f	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	/replace(2859, A)	
FT	/*tag= g	
FT	/standard_name= "Single nucleotide polymorphism"	
FT	/replace(3420, C)	
FT	/*tag= h	
FT	/standard_name= "Single nucleotide polymorphism"	
XX		
XX	W0200240683-A2.	
PN		
XX		
PD	23-MAY-2002..	
XX		
XX	22-OCT-2001; 2001WO-US032616.	
XX		
PR	14-NOV-2000; 2000US-00711124.	
PR	17-MAY-2001; 2001US-00858664.	
XX		
XX	(PEKE) PE CORP NY.	
PA		
XX		
PI	Wei M, Ketchum K, Di Francesco V, Beasley EM;	
XX		
DR	WPI; 2002-500223/53.	
DR	P-PSDB; AAO15372.	
XX		
PT	New kinase proteins related to myosin light chain kinase subfamily and encoding polynucleotide, useful for diagnosing, treating disease or condition mediated by the kinase protein and for identifying modulators.	
PT		
PT		
XX		
PS	Claim 23; Fig 3; 96pp; English.	
XX		
CC	The invention comprises the amino acid and coding sequences (located on chromosome 1) of a human kinase protein that is related to the myosin light chain kinase subfamily. The human kinase DNA and protein sequences of the invention are useful for identifying agents that modulate the activity of the human kinase protein. Kinase-modulating agents are useful for treating a disease or condition mediated by a human kinase protein. The human kinase DNA sequences can be used to produce transgenic animals which are useful for studying the function of kinase proteins and identifying/evaluating modulators of kinase protein activity. The present DNA sequence encodes the human kinase protein of the invention	
XX		
SQ	Sequence 5207 BP; 988 A; 1758 C; 1612 G; 849 T; 0 U; 0 Other;	
	Query Match 99.8%; Score 999.4; DB 6; Length 5207;	
	Best Local Similarity 99.9%; Pred. No. 4.3e-197;	
	Matches 1000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 GGCAGTCAGTGACACTGGCTGCCAGTGTGACGCCAGCCAGCTGCCAGGCCACCTGG 60	
Db	3628 GGCAGTCAGTGACACTGGCTGCCAGTGTGACGCCAGCCAGCTGCCAGGCCACCTGG 3687	
QY	61 AGCAAGACGGAGCCCCCTTGGAGAGCAGCAGCGTGTCTCATCTTGCACCCCTCAAG 120	

PD 26-OCT-2000.
XX
PF 11-APR-2000; 2000WO-US009488.
XX
PR 16-APR-1999; 99US-0129553P.
XX (SCIO-) SCIOS INC.
PA
XX Zeng W, Stanton L, Kong H;
PI WPI; 2001-007013/01.
XX P-PSDB; AAB30569.
DR
XX Novel h19G5 polypeptides capable of regulating signal transduction and
PT exhibiting kinase activity useful for identifying antibodies to treat
PT cardiac diseases, and additional mediators of signal transduction.
XX
XX Claim 4; Page 65-68; 81pp; English.
PS
XX The present sequence encodes a splice variant of human in signal
CC transduction polypeptide. The polypeptide is designated H19G5. The
CC protein is capable of regulating signal transduction and exhibits kinase
CC activity. The H19G5 transcript is expressed in the heart. H19G5
CC polypeptides and polynucleotides are useful for preventing or treating a
CC cardiac disease, such as congestive heart failure, dilated congestive
CC cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy,
CC mitral valve disease, aortic valve disease or tricuspid valve disease,
CC angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary,
CC arterial or renovascular hypertension, arteriosclerosis, atherosclerosis
CC and cardiac tumours in humans. The polypeptide is also useful for
CC detecting the expression of a protein capable of regulating signal
CC transduction or the expression of a protein capable of acting as a donor
CC or acceptor molecule of a phosphate group. The monoclonal antibodies can
CC be used as probes for detecting discrete antigens expressed by tissue or
CC cell samples, and therefore used in humans for localization and
CC monitoring of microbial infection
XX
SQ Sequence 7928 BP; 1535 A; 2644 C; 2488 G; 1261 T; 0 U; 0 Other;

Query Match 99.8%; Score 999.4; DB 5; Length 7928;
Best Local Similarity 99.9%; Pred. No. 4.5e-197;
Matches 1000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCCAGTCAGTGACACTGCGCTCCAGGTGTGAGCCGACGAGCTGCCAGGCCACCTGG 60
DB 6372 GCCCAGTCAGTGACACTGCGCTCCAGGTGTGAGCCGACGAGCTGCCAGGCCACCTGG 6431
QY 61 AGCAAAGACGGAGCCCTGAGAGAGCAGCAGCGGTGCTCATCTCTGCCACCTCAAG 120
DB 6432 AGCAAAGACGGAGCCCTGAGAGAGCAGCAGCGGTGCTCATCTCTGCCACCTCAAG 6491
QY 121 AACTTCCAGCTTCTGACCATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 180
DB 6492 AACTTCCAGCTTCTGACCATCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 6551
QY 181 AGCTGAGCAATGCGCTGGGACAGTACACACAGCGGGGTCTCCGGAAGCAGAGCGC 240
DB 6552 AGCTGAGCAATGCGCTGGGACAGTACACACAGCGGGGTCTCCGGAAGCAGAGCGC 6611
QY 241 CCCTCATCTTCCGCAATGCCGATATCGGGAGGTGTACCGGATGGGGTGTCTGTGTC 300
DB 6612 CCCTCATCTTCCGCAATGCCGATATCGGGAGGTGTACCGGATGGGGTGTCTGTGTC 6671
QY 301 TGGAAAGCCGCTGGAATCTTACCGCCCTGTGACCTACATTTGTGACGTGACGCTTAGAAGGC 360
DB 6672 TGGAAAGCCGCTGGAATCTTACCGCCCTGTGACCTACATTTGTGACGTGACGCTTAGAAGGC 6731
QY 361 GGCAGCTGACACACACTGCGCTCCGACATCTTTTGACTGTCTGTCTGTCTGTCTGTCTGTCT 420
DB 6732 GGCAGCTGACACACACTGCGCTCCGACATCTTTTGACTGTCTGTCTGTCTGTCTGTCTGTCT 6791
QY 421 TCCCGGGGTGGCAGCTTACACCTTCCGACGCGCATGTGTGACGAGGAGGATGGTCCC 480

Db 6792 TCCCGGGGTGGCAGCTTACACCTTCCGACGCGCATGTGTGACGAGGAGGATGGGTCCC 6851
QY 481 TACAGACGCCCTCGAGCAAGTCTCTCTGGAGGGCCGACGCCACTGGCTTCTGAGGAG 540
Db 6852 TACAGACGCCCTCGAGCAAGTCTCTCTGGAGGGCCGACGCCACTGGCTTCTGAGGAG 6911
QY 541 GAGAGCCAGGGGGCGGTGAGCCCAACCCCTGCCACGACACAAAGACCTTCGATTCAGACA 600
Db 6912 GAGAGCCAGGGGGCGGTGAGCCCAACCCCTGCCACGACACAAAGACCTTCGATTCAGACA 6971
QY 601 CAGATCCAGAGGGGGCGGTTCAGCGTGTGCGGCAATGCTGGGAGGAGGCCAGCGGGCGG 660
Db 6972 CAGATCCAGAGGGGGCGGTTCAGCGTGTGCGGCAATGCTGGGAGGAGGCCAGCGGGCGG 7031
QY 661 GCGTGGCGGGCAAGATCATCCCTTACCAACCCCAAGGAGCAAGACAGTGTGCGCGAA 720
Db 7032 GCGTGGCGGGCAAGATCATCCCTTACCAACCCCAAGGAGCAAGACAGTGTGCGCGAA 7091
QY 721 TACGAGGCCCTCAAGGGCTGCGCCACCCGACCTGGCCGACGCTGCACGCGACCTACCTC 780
Db 7092 TACGAGGCCCTCAAGGGCTGCGCCACCCGACCTGGCCGACGCTGCACGCGACCTACCTC 7151
QY 781 AGCCCCCGGACCTGGTGTCTCATCTTGAGGTGTCTCTGGGCCCGAGCTGTCTCCCTGTC 840
Db 7152 AGCCCCCGGACCTGGTGTCTCATCTTGAGGTGTCTCTGGGCCCGAGCTGTCTCCCTGTC 7211
QY 841 CTGGCCGAGAGGGGCTCTCTACTCAGATCCGAGGTGAAGGACTACTCTGTGGCAGATGTTG 900
Db 7212 CTGGCCGAGAGGGGCTCTCTACTCAGATCCGAGGTGAAGGACTACTCTGTGGCAGATGTTG 7271
QY 901 AGTGGCAGCCAGTACCTGCACACAGCAGCATCTCTGACCTGGACCTGAGGTCCGAGAC 960
Db 7272 AGTGGCAGCCAGTACCTGCACACAGCAGCATCTCTGACCTGGACCTGAGGTCCGAGAC 7331
QY 961 ATGATCATCACCAGTAATACAACTGCTCAAGGTGCTGGACCT 1001
Db 7332 ATGATCATCACCAGTAATACAACTGCTCAAGGTGCTGGACCT 7372

RESULT 11
ADR41256
ID ADR41256 standard; cDNA; 876 BP.
XX
AC ADR41256;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human CD-like molecule HAHB26 cDNA, SEQ ID NO:55.
XX
KW Human; CD-like molecule; cluster of differentiation; diagnosis;
KW prevention; immune disorder; immunodeficiency; autoimmune disorder;
KW blood-related disorder; haematological disorder; haemostatic disorder;
KW thrombolytic disorder; hyperproliferative disorder; cancer; tumour;
KW apoptotic disorder; cardiovascular disorder; respiratory disorder;
KW angiogenic disorder; neovascularisation; neurological disorder;
KW endocrine disorder; reproductive system disorder; infectious disease;
KW gastrointestinal disorder; drug screening; tissue regeneration;
KW chemotaxis; gene therapy; antibody therapy; drug targeting;
KW chromosome mapping; forensic analysis; immunophenotyping; cytostatic;
KW haemostatic; tranquiliser; vulnery; antiinflammatory; nephrotropic;
KW cardiant; aniallergic; anti-HIV; antirheumatic; antiarthritic;
KW antipruritic; immunosuppressive; vasotropic; nootropic; neuroprotective;
KW antithyroid; thymolymetic; gynaecological; virucide; hepatotropic;
KW antibacterial; dermatological; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200226930-A2.
XX
PD 04-APR-2002.
XX
PF 25-SEP-2001; 2001WO-US029838.
XX

PR 26-SEP-2000; 2000US-0235484P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Birse CE;
XX
XX WPI; 2002-405050/43.
DR P-PSDB; ADR41432.
XX
XX Novel polynucleotides and polypeptides useful for treating, preventing or
PT ameliorating cardiovascular, renal, neurovascular, and autoimmune
PT disorders.
XX
XX Claim 4; SEQ ID NO 55; 1243bp; English.
PS
XX
XX The invention relates to 167 novel human CD (cluster of differentiation)-
CC like molecules (ADR41388-ADR41563) and to cDNAs encoding them (seqid:11)-
XX
XX Sequence 876 BP; 178 A; 279 C; 268 G; 141 T; 0 U; 10 Other;
SQ
Query Match 38.5%; Score 385; DB 7; Length 876;
Best Local Similarity 100.0%; Pred. No. 3.9e-70;
Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 617 GCTTCAGCGTGTGGCGCAATCTGGGAGAGGCGGCGGCGCTGGCCGCCAAGA 676
Db 125 GCTTCAGCGTGTGGCGCAATCTGGGAGAGGCGGCGGCGCTGGCCGCCAAGA 184
QY 677 TCATCCCTTACCACCCCAAGGACAGACAGCAGTGTGGCGAATACAGGCCCTCAAGG 736
Db 185 TCATCCCTTACCACCCCAAGGACAGACAGCAGTGTGGCGAATACAGGCCCTCAAGG 244
QY 737 GCTTCAGCGCGCCAGCTGGCCCGCAGCTGACGCGAGCTTACTCAGCCCCCGGCACCTGG 796
Db 245 GCTTCAGCGCGCCAGCTGGCCCGCAGCTGACGCGAGCTTACTCAGCCCCCGGCACCTGG 304
QY 797 TGCTCATCTTGAGCTGTGCTCTGGGCGGAGTGTCTCCCTGCTGGCCCGAGAGGGCCT 856
Db 305 TGCTCATCTTGAGCTGTGCTCTGGGCGGAGTGTCTCCCTGCTGGCCCGAGAGGGCCT 364
QY 857 CCTACTCAGAAATCCGAGGTGAAGGACTACCTGTGGCAGATGTTGAGTGCCACCCAGTACC 916
Db 365 CCTACTCAGAAATCCGAGGTGAAGGACTACCTGTGGCAGATGTTGAGTGCCACCCAGTACC 424
QY 917 TGCAACACGACATCTCTGCACCTGGACCTGAGTCCGAGAACATGATCATCATCCGAAAT 976
Db 425 TGCAACACGACATCTCTGCACCTGGACCTGAGTCCGAGAACATGATCATCATCCGAAAT 484
QY 977 ACACCTGCTCAAGTCTGGACCT 1001
Db 485 ACAACCTGCTCAAGTCTGGACCT 509
RESULT 12
AAC25654
ID AAC25654 standard; cDNA; 182 BP.
XX
XX AAC25654;
XX
XX 06-OCT-2000 (first entry)
XX
XX Human secreted protein 5' EST, SEQ ID NO: 29729.
XX
XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX
XX Homo sapiens.
XX
XX EP1033401-A2.
XX
XX 06-SEP-2000.
XX
XX 21-FEB-2000; 2000EP-00200610.

XX 26-FEB-1999; 99US-0122487P.
PR (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Duclert A, Giordano J;
PI WPI; 2000-500381/45.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures.
XX
XX Claim 1; SEQ ID NO 29729; 71pp + Sequence Listing; English.
PS
XX
XX The present sequence is one of a large number of 5' ESTs derived from
CC mRNAs encoding secreted proteins. No ORF has yet been conclusively
CC identified within the present sequence. The 5' ESTs were prepared from
CC total human RNAs or polyA+ RNAs derived from 30 different tissues. EST
CC sequences usually correspond mainly to the 3' untranslated region (UTR)
CC of the mRNA because they are often obtained from oligo-dT primed cDNA
CC libraries. Such ESTs are not well suited for isolating cDNA sequences
CC derived from the 5' ends of mRNAs and even in those cases where longer
CC cDNA sequences have been obtained, the full 5' UTR is rarely included. 5'
CC ESTs are derived from mRNAs with intact 5' ends and can therefore be used
CC to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in
CC diagnostic, forensic, gene therapy and chromosome mapping procedures.
CC They are used to obtain upstream regulatory sequences and to design
CC expression and secretion vectors
XX
SQ Sequence 182 BP; 34 A; 69 C; 57 G; 19 T; 0 U; 3 Other;
Query Match 16.5%; Score 164.8; DB 3; Length 182;
Best Local Similarity 97.8%; Pred. No. 1.1e-24;
Matches 174; Conservative 3; Mismatches 0; Indels 1; Gaps 1;
QY 613 GGCGCTTTCAGCTGTGGCGCAATGCTGGGAGAGGCCACGCGGCGGCGCTGGCGGCC 672
Db 1 GGCGCTTTCAGCTGTGGCGCAATGCTGGGAGAGGCCACGCGGCGGCGCTGGCGGCC 60
QY 673 AAGATCATTCCTTACCAACCCCAAGGACAGACAGCAGTGTGGCGGAATACGAGGCCCTC 732
Db 61 AAGATCATTCCTTACCAACCCCAAGGACAGACAGCAGTGTGGCGGAATACGAGGCCCTC 120
QY 733 AAGGCGCTGGCGCCACCGCACCTGGCCAGCTGCA-CGCAAGCTTACTCAGCCCCCGG 789
Db 121 AAGGCGCTGGCGCCACCGCACCTGGCCAGCTGCA-CGCAAGCTTACTCAGCCCCCGG 178
RESULT 13
ADE47671
ID ADE47671 standard; DNA; 9930 BP.
XX
XX ADE47671;
XX
XX 29-JAN-2004 (first entry)
XX
XX Human NOV14a gene SEQ ID NO:33.
XX
XX ds; gene; human; cardiant; antiarteriosclerotic; hypotensive;
KW immunosuppressive; dermatological; anorectic; cytostatic; antidiabetic;
KW haemostatic; anti-HIV; antiasthmatic; antibacterial; virucide;
KW neuroprotective; nootropic; antiparkinsonian; antilipaemic; gene therapy;
KW vaccine.
XX
XX Homo sapiens.
XX
XX WO2003076642-A2.
XX
XX 18-SEP-2003.
XX
XX 02-AUG-2002; 2002WO-US024459.
XX

KW cancer; autoimmune disorder; stroke; inflammatory bowel disease;
KW inflammatory pelvic disease; multiple sclerosis; psoriasis; ss.
XX Homo sapiens.

OS WO200073469-A2.

PN PD

PD 07-DEC-2000.

XX 26-MAY-2000; 2000WO-US014842.

XX 28-MAY-1999; 99US-0136503P.

XX (SUGE-) SUGEN INC.

XX Plowman GD, Martinez R, Whyte D, Sudersanam S;

XX WPI; 2001-032161/04.

XX P-PSDB; AAB65635.

XX Nucleic acids encoding kinase polypeptides, useful for diagnosing and

PT treating immune-related diseases and disorders, cardiovascular disease,

PT neurodegenerative diseases and/or cancers.

XX Example 1; Fig 2; 310pp; English.

XX The present sequence encodes a novel protein kinase. The nucleic acids

CC and the protein kinases they encode may be used in the treatment and

CC diagnosis of diseases associated with inappropriate kinase expression

CC such as immune-related diseases and disorders, cardiovascular disease,

CC neurodegenerative diseases and/or cancers. The nucleic acids and

CC complementary sequences may also be used as DNA probes in diagnostic

CC assays. The kinase polypeptides may be used as antigens in the production

CC of antibodies of kinase expression and activity. Anti-kinase antibodies

CC and kinase antagonists may also be used to down regulate kinase

CC expression and activity. Diseases related to kinase expression and

CC activity include rheumatoid arthritis, atherosclerosis, autoimmune

CC disorders, complications of organ transplantation, myocardial infarction,

CC immune disorders, cardiomyopathies, strokes, renal failure, oxidative-

CC stress related disorders, chronic inflammatory bowel disease, chronic

CC inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis,

CC psoriasis, rhinitis, autoimmunity, diabetes, cancers and reproductive

CC disorders

XX

SQ Sequence 7710 BP; 1446 A; 2534 C; 2447 G; 1283 T; 0 U; 0 Other;

Query Match 15.0%; Score 150; DB 4; Length 7710;
Best Local Similarity 59.7%; Pred. No. 2e-21;
Matches 252; Conservative 0; Mismatches 170; Indels 0; Gaps 0;

QY 580 AAGACCTTGGCATTCAGACACAGATCCAGAGGGGCGCTTCAGCGTGGTGGCGCAATGC 639
DB |||||
QY 640 TGGGAGAGCCAGCGGGGCGCTGGCGCCAGATCATCCCTACCAACCCCAAGGAC 699
DB |||||
QY 6007 CGGAGAGATGCCACGGGGCGAAGCTTCGTGGCCAAAGATCGTGCCCTATGCTGCCGAGGCG 6066
DB |||||
QY 700 AAGACGAGTGTCTCGCGAATACAGAGCCCTCAAGGGCGCTCGCCACCCGACCTGGCC 759
DB |||||
QY 6067 AAGCGCGGGTCTCGAGAGTACAGAGTGTGGCGACCTGACACAGCGGATCATG 6126
DB |||||
QY 760 CAGCTGCACGACGCTACTCAGCCCCCGGCACTGGTGTCTCATCTTGGAGCTGTGCTCT 819
DB |||||
QY 6127 TCCCTGCACGAGGCGCTACATACCCCTCGGTACCTCGTCTCATTTGCTGAGAGCTGTGGC 6186
DB |||||
QY 820 GGGCCCGAGCTGCTCCCTGCTGGCGAGAGGGGCTCTACTCAGATCCGAGGTGAAG 879
DB |||||
QY 6187 AACCGGGAATCTCTGTGGGCTCAGTGACAGGTTCCGGTATTTCTGAGGATGACGTGGCC 6246
DB |||||
QY 880 GACTACCTGTGGCAGATGTGAGTGCCACCCAGCTACCTGCACCAACAGCAGCATCTCGCAC 939
DB |||||
QY 6247 ACTTACATGGTGACGCTGTACAAGGCTTGAGTACCTCCACGGCCACACGCTGCTCCAC 6306
DB |||||

QY 940 CTGACCTGAGGTCGAGAACATGATCATCAACGATACACCTGCTCAAGGTGCTGAC 999
DB |||||
QY 6307 CTAGACATCAAGCCAGACAACTGCTGCTGGCCCTGACAATGCGCTCAAGATTGTGAC 6366
DB |||||
QY 1000 CT 1001
DB 6367 TT 6368

Search completed: March 20, 2005, 18:54:57
Job time : 586.044 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:15:10 ; Search time 581.044 Seconds
(without alignments)
10198.307 Million cell updates/sec

Title: US-10-077-130-4_COPY_3120_4120

Perfect score: 1001
Sequence: 1 aaagatgggaagaagctgag.....tcagagcccaagcggtgtt 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04:.*
1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002as:.*
7: geneseqn2002bs:.*
8: geneseqn2003as:.*
9: geneseqn2003bs:.*
10: geneseqn2003cs:.*
11: geneseqn2003ds:.*
12: geneseqn2004as:.*
13: geneseqn2004bs:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	24120	8 ABX11642	Abx11642 Human ser
2	965	96.4	20489	12 ADQ22881	Adq22881 Human sof
3	248.4	24.8	2768	8 ABX71198	Abx71198 Novel hum
4	218.6	21.8	1251	10 ADC30509	Adc30509 Human nov
5	218.6	21.8	2488	4 AAH18453	Aah18453 Human cdn
6	218	21.8	3956	13 ADR08183	Adr08183 Full leng
7	140.2	14.0	2737	4 AA159251	Aa159251 Human pol
8	140.2	14.0	5378	13 ADR67197	Adr67197 Human bla
9	140.2	14.0	5382	10 ADD14722	Add14722 Human src
10	130.2	13.0	1458	12 ADL24314	Adl24314 AW755252-
11	123.8	12.3	3999	4 AA161037	Aa161037 Human pol
12	123.2	12.3	1645	5 Aaf24162	Aaf24162 Human sec
13	123.2	12.3	1948	2 ABV74346	Abv74346 Human il-
14	123.2	12.3	1949	2 ABV74347	Abv74347 Human il-
15	123.2	12.3	2170	6 ABQ54970	Abq54970 Human ova
16	123	12.3	2184	4 AAH18256	Aah18256 Human cdn
17	107	10.7	2559	8 ACC46215	Acc46215 Human dit
18	104	10.4	724	4 AAH08421	Aah08421 Human cdn
19	97.4	9.7	14061	6 ABV99363	Abv99363 Human nov
20	97.4	9.7	14109	6 ABV99362	Abv99362 Human nov

21	95.8	9.6	2534	11 ADM03261	Adm03261 Human cdn
22	95.8	9.6	7564	8 ABZ24581	Abz24581 Human cel
23	88.4	8.8	2254	3 AAC93403	Aac93403 Human sec
24	88.4	8.8	2254	5 AAF24183	Aaf24183 Human sec
25	71	7.1	572	5 AAS42508	Aas42508 Human cdn
26	68	6.8	716	6 ABK35726	Abk35726 cDNA sequ
27	66.8	6.7	9591	4 AAS28906	Aas28906 Human inm
28	66.8	6.7	9591	4 AAL03355	Aal03355 Human rep
29	66.8	6.7	9591	10 ADB31747	Adb31747 Human nov
30	66.8	6.7	12415	4 AAS28907	Aas28907 Human inm
31	66.8	6.7	12415	4 AAL03356	Aal03356 Human rep
32	66.8	6.7	12415	10 ADB31748	Adb31748 Human nov
33	66.8	6.7	20565	4 AAK89083	Aak89083 Human dig
34	66.8	6.7	20565	4 AAS28908	Aas28908 Human inm
35	66.8	6.7	20565	4 AAL03357	Aal03357 Human rep
36	66.8	6.7	20565	10 ADB31749	Adb31749 Human nov
37	62.2	6.2	30191	12 ADQ97654	Adq97654 Mouse can
38	62	6.2	31595	10 ADF81661	Adf81661 Leukaemia
39	62	6.2	81940	4 AAS05390	Aas05390 Human tit
40	62	6.2	81940	6 ABK64829	Abk64829 Human ben
41	62	6.2	81940	12 ADQ17315	Adq17315 Human sof
42	62	6.2	93801	9 ABX13540	Abx13540 Human RGS
43	62	6.2	103052	13 ADQ89963	Adq89963 Antagonis
44	61.6	6.2	1266	6 ABK35725	Abk35725 cDNA sequ
45	61.6	6.2	20448	4 ABL19989	Ab119989 Drosophil

ALIGNMENTS

RESULT 1
ABX11642
ID ABX11642 standard; cDNA; 24120 BP.
XX
AC ABX11642;
XX
DT 09-MAY-2003 (first entry)
XX
DE Human serine/threonine or protein kinase 12599, cDNA.
XX
KW Human; ss; Gene; serine/threonine kinase; protein kinase; 12599;
KW cardiovascular disease; heart failure; myocardial infarction;
KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma;
KW blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
KW haemolytic anaemia; cellular proliferative disorder; cancer;
KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
KW psoriasis; inflammatory bowel disease; rheumatoid arthritis;
KW multiple sclerosis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..71
FT CDS /tag= a
FT /tag= b
FT /tag= c
FT 3'UTR /product= "Kinase 12599"
FT /note= "This CDS is specifically claimed in claim 2"
FT 23979..24120
FT /tag= c
XX
US2002169742-A1.
XX
14-NOV-2002.
XX
15-FEB-2002; 2002US-00077130.
XX
15-FEB-2001; 2001US-0269201P.
XX
(MILL-) MILLENNIUM PHARM INC.
XX
Kapeller-Libermann R, Acton SL;
XX

DR WPI; 2003-298729/29.
DR P-PSDB; ABG76187.
XX
PT Novel isolated human protein kinase, designated 59079 or 12599
PT polypeptide, useful as diagnostic and therapeutic agents for preventing
PT cardiovascular diseases, proliferative disorders, and protein kinase
PT disorders.
XX
PS Claim 2; Page 58-84; 119pp; English.
XX
CC The invention relates to an isolated human serine/threonine or protein
CC kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule
CC comprising at least 85% identity to the nucleic acids appearing as
CC ABX11641 and ABX11642 or their complement, a naturally occurring variant
CC of the kinases or their fragments. Also included are a non-human host
CC cell containing the nucleic acids, an antibody specific for the proteins,
CC identifying a compound which binds to the kinase (by contacting the
CC kinase or a cell expressing the kinase with a test compound and
CC determining whether the kinase binds to the test compound) and modulating
CC the activity of kinase using the identified compound. The kinases and
CC their encoding nucleic acids are useful as diagnostic and therapeutic
CC agents for preventing a disease or condition associated with an aberrant
CC or unwanted 59079 or 12599 activity in a subject, including
CC cardiovascular diseases such as heart failure, and myocardial infarction;
CC disorders involving blood vessels such as atherosclerosis, and Kaposi's
CC sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia,
CC Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders
CC such as cancer; and protein kinase disorders such as autoimmune
CC disorders, diabetes mellitus, psoriasis, inflammatory bowel disease,
CC rheumatoid arthritis, and multiple sclerosis (many examples of diseases
CC and disorders are included in the specification). The kinases, their
CC encoding nucleic acids and antibodies are useful in screening assays,
CC detection assays (e.g. forensic biology), and predictive medicine (e.g.
CC diagnostic assays, prognostic assays, and monitoring clinical trials and
CC pharmacogenomics). The kinases and their encoding nucleic acids are
CC useful as query sequences to perform a search against public databases to
CC identify other family members or related sequences. The present sequence
CC encodes the kinase 12599
XX
SQ Sequence 24120 BP; 4719 A; 7265 C; 8278 G; 3858 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 8; Length 24120;
Best Local Similarity 100.0%; Pred. No. 2.2e-226;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAAGATGGGAGAGAGCTGAGCTCCAGCTTGAAGTGCATGTAGAGGCCAAAGGCTGCAGA 60
DB |||||||
QY 61 CGAGGCTGGTGGTGCAGAGCGAGCAAGACGGATGCCGGGACTACAGCTGCGAGGCC 120
DB |||||||
QY 121 AGGGCCACAGAGGCTCTCCCTCGCTGCACATCACAGAGCCCAAGATGATGTTGCAAG 180
DB |||||||
QY 181 GAGCAGTCAGTCATAATAGGTGACGCTGAGCGGGGGCCAGTGCCATGCTGAGCTGT 240
DB |||||||
QY 241 GAGTGGCCCGAGCCCGACAGCGAGGTGACGTGGTACAAGATGGGAAGAGCTGAGCTCC 300
DB |||||||
QY 301 AGCTCAAAAGTGGGATCAGAGGTCAAGGGTGCACAGGAGGCTGGTCTGCCACAGCG 360
DB |||||||
QY 3420 AGCTCAAAAGTGGGATCAGAGGTCAAGGGTGCACAGGAGGCTGGTCTGCCACAGCG 3479
QY 361 GGCAAGCAGATGCTGGGAGTACAGCTGTGAGGCTGGGGGCCAGAGAGTCTCTCCAC 420
DB |||||||
QY 3480 GGCAAGCAGATGCTGGGAGTACAGCTGTGAGGCTGGGGGCCAGAGAGTCTCTCCAC 3539

QY 421 CTGCACATCACAGAGCCCAAGGGGGTGTTCGGAAGGAGCAGTCAGTGCATAATGAGGTG 480
DB |||||||
QY 3540 CTGCACATCACAGAGCCCAAGGGGGTGTTCGGAAGGAGCAGTCAGTGCATAATGAGGTG 3599
DB |||||||
QY 481 CAGGCTGAGGCGGGGACCACTGCTAGCTGTGAGGTGGCCAGCCGCCAGCGAG 540
DB |||||||
QY 3600 CAGGCTGAGGCGGGGACCACTGCTAGCTGTGAGGTGGCCAGCCGCCAGCGAG 3659
DB |||||||
QY 541 GTGACGTGGTCAAAAGGAGCGGAAGAGCTGAGCTCCAGCTCAAAGTACGATGAGGTG 600
DB |||||||
QY 3660 GTGACGTGGTCAAAAGGAGCGGAAGAGCTGAGCTCCAGCTCAAAGTACGATGAGGTG 3719
DB |||||||
QY 601 AAGGGCTGCACACGAAGGCTGGTAGTCAGCAGGTGGGCAAGAGCAGATGCTGGGGAGTAC 660
DB |||||||
QY 3720 AAGGGCTGCACACGAAGGCTGGTAGTCAGCAGGTGGGCAAGAGCAGATGCTGGGGAGTAC 3779
DB |||||||
QY 661 AGCTGCGAGGCTGGGGGCCAGAGAGTCTCTTTCACTGCACATCACAGAGCCCAAGGCA 720
DB |||||||
QY 3780 AGCTGCGAGGCTGGGGGCCAGAGAGTCTCTTTCACTGCACATCACAGAGCCCAAGGCA 3839
DB |||||||
QY 721 GTGTTTGCAGAGGAGCAGTTGGTGCATAATGAGGTGCGGACTGAGGCGAGGGGCCAGTGCC 780
DB |||||||
QY 3840 GTGTTTGCAGAGGAGCAGTTGGTGCATAATGAGGTGCGGACTGAGGCGAGGGGCCAGTGCC 3899
DB |||||||
QY 781 ACACGTGAGCTGTGAGGTGGGCCAGGCCAGACAGAGGTGACGTGTTACAAGATGGGAAG 840
DB |||||||
QY 3900 ACACGTGAGCTGTGAGGTGGGCCAGGCCAGACAGAGGTGACGTGTTACAAGATGGGAAG 3959
DB |||||||
QY 841 AAGCTGAGCTCCAGTTTCGAAAGTCCGATAGAGCTCGGGCTGCATGCGGACAGCTGGTG 900
DB |||||||
QY 3960 AAGCTGAGCTCCAGTTTCGAAAGTCCGATAGAGCTCGGGCTGCATGCGGACAGCTGGTG 4019
DB |||||||
QY 901 GTGCAGCAGGCGAGGCCAGGAGATGCTGGGAGGTACACCTGTGAGGCTGGGGGCCAGCGG 960
DB |||||||
QY 4020 GTGCAGCAGGCGAGGCCAGGAGATGCTGGGAGGTACACCTGTGAGGCTGGGGGCCAGCGG 4079
DB |||||||
QY 961 CTCTCTTCCACCTGGATGTTTCAGAGCCCCAAGCGGTGTT 1001
DB |||||||
QY 4080 CTCTCTTCCACCTGGATGTTTCAGAGCCCCAAGCGGTGTT 4120
DB |||||||

RESULT 2
ADQ22881
ID ADQ22881 standard; DNA; 20489 BP.
AC AC
XX AC ADQ22881;
XX DT 26-AUG-2004 (first entry)
XX DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5701.
XX DE soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
XX KW ds.
XX OS Homo sapiens.
XX PN WO2004048938-A2.
XX PD 10-JUN-2004.
XX PF 26-NOV-2003; 2003WO-US038193.
XX PR 26-NOV-2002; 2002US-0429739P.
XX PA (PROT-) PROTEIN DESIGN LABS INC.
XX PI Aziz N, Ginsburg WM, Zlotnik A;
XX WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue

PT sarcoma.
 XX Example 2; SEQ ID NO 5701; 210pp; English.
 PS
 XX The invention relates to a novel method for detecting soft tissue sarcoma
 CC which comprises obtaining a first soft tissue sample from an individual
 CC and a normal soft tissue sample from the same or different individual,
 CC determining the expression of a gene in both samples and comparing the
 CC expression of the gene in both soft tissue samples, where a higher level
 CC of protein expression in the first soft tissue sample indicates the
 CC presence of soft tissue sarcoma. The method of the invention has
 CC cytosstatic applications and may be useful for detecting soft tissue
 CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
 CC acid sequences may be useful in diagnostic and screening applications.
 CC The current sequence is that of a human soft tissue sarcoma-upregulated
 CC DNA of the invention. The current sequence is not shown within the
 CC specification per se but was submitted in CD format by the inventor.
 XX
 SQ Sequence 20489 BP; 3917 A; 5815 C; 6983 G; 3225 T; 0 U; 549 Other;

Query Match 96.4%; Score 965; DB 12; Length 20489;
 Best Local Similarity 96.4%; Pred. No. 6.9e-218;
 Matches 965; Conservative 0; Mismatches 36; Indels 0; Gaps 0;

QY 1 AAAGATGGGAAGAGCTGAGCTCCAGCTTGAAGTGCATGTAGAGCCAAAGGCTGCAGA 60
 DB 3120 AAAGATGGGAAGAGCTGAGCTCCAGCTTGAAGTGCATGTAGAGCCAAAGGCTGCAGA 3179
 QY 61 CGAGGCTGGTGGTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 120
 DB 3180 CGAGGCTGGTGGTGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3239
 QY 121 AGGGCCAGAGGCTTCTTCCGCTGCACATCAAGAGCCCAAGATGATGTTTCAAG 180
 DB 3240 AGGGCCAGAGGCTTCTTCCGCTGCACATCAAGAGCCCAAGATGATGTTTCAAG 3299
 QY 181 GAGCAGTCAGTGCATTAATGAGTGCAGCTGAGCGGGGGGCGAGTGCATGCTGAGCTGT 240
 DB 3300 GAGCAGTCAGTGCATTAATGAGTGCAGCGNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3359
 QY 241 GAGTGGCCAGGCCCCAGACGAGGTGACGTGGTCAAGAGGTGGGAAGAGCTGAGCTCC 300
 DB 3360 NNNNNGGCCAGGCCCCAGACGAGGTGACGTGGTCAAGAGGTGGGAAGAGCTGAGCTCC 3419
 QY 301 AGCTCAAAAGTGGGATGAGGTCAAGAGGTGCAACGAGGCTGGTGTGCAACAGCGG 360
 DB 3420 AGCTCAAAAGTGGGATGAGGTCAAGAGGTGCAACGAGGCTGGTGTGCAACAGCGG 3479
 QY 361 GGCAAGCAGATGCTGGGAGTACAGCTGTGAGGCTGGGGCCAGAGAGTCTCTTCCAC 420
 DB 3480 GGCAAGCAGATGCTGGGAGTACAGCTGTGAGGCTGGGGCCAGAGAGTCTCTTCCAC 3539
 QY 421 CTGCATCATCAGAGCCCAAGGGGTGTTTGGGAAGGAGCAGTCAAGTGCATTAATGAGTGT 480
 DB 3540 CTGCATCATCAGAGCCCAAGGGGTGTTTGGGAAGGAGCAGTCAAGTGCATTAATGAGTGT 3599
 QY 481 CAGGCTGAGGGGGGACCACTGCCATGCTGAGCTGTGAGGCTGGGGCCAGAGAGTGTG 540
 DB 3600 CAGGCTGAGGGGGGACCACTGCCATGCTGAGCTGTGAGGCTGGGGCCAGAGAGTGTG 3659
 QY 541 CTGAGCTGTTACAGAGGGGAGAGTGCAGCTCCAGCTCAAAAGTACGATGAGTGT 600
 DB 3660 GTGAGCTGGTACAGAGGGGAGAGTGCAGCTCCAGCTCAAAAGTACGATGAGTGT 3719
 QY 601 AAGGCTGCACACGAGGCTGGTGTGAGTGCAGAGGTGGGCAAGCAGATGCTGGGAGTAC 660
 DB 3720 AAGGCTGCACACGAGGCTGGTGTGAGTGCAGAGGTGGGCAAGCAGATGCTGGGAGTAC 3779
 QY 661 AGCTCGAGGCTGGGGGCGAGAGTGTCTTTCAACTGCACATCACAGAGCCCAAGGCA 720
 DB 3780 AGCTCGAGGCTGGGGGCGAGAGTGTCTTTCAACTGCACATCACAGAGCCCAAGGCA 3839
 QY 721 GTGTTTCCACAGGACAGTGTGTCATTAATGAGGTGCGGACTGAGGCGAGGGGCCAGTGCC 780

DB 3840 GTGTTTCCACAGGACAGTGTGTCATTAATGAGGTGCGGACTGAGGCGAGGGCCAGTGCC 3899
 QY 781 ACATGAGCTGTGAGGTGGCCAGGCCCAGACAGAGGTGACGTGTGTAACAAGATGGGAAG 840
 DB 3900 ACATGAGCTGTGAGGTGGCCAGGCCCAGACAGAGGTGACGTGTGTAACAAGATGGGAAG 3959
 QY 841 AAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGCTCGGGCTGCATCGGCAGCTGGTG 900
 DB 3960 AAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGCTCGGGCTGCATCGGCAGCTGGTG 4019
 QY 901 GTGCAGCAGGCGAGCCAGGAGATGCTGGGAGTACACCTGTGAGGCTGGGGGCCAGCGG 960
 DB 4020 GTGCAGCAGGCGAGCCAGGAGATGCTGGGAGTACACCTGTGAGGCTGGGGGCCAGCGG 4079
 QY 961 CTCCTCTTCCACCTGGATGTTTCAGAGCCCAAGCGGTGT 1001
 DB 4080 CTCCTCTTCCACCTGGATGTTTCAGAGCCCAAGCGGTGT 4120

RESULT 3

ABX71198

ID ABX71198 standard; cDNA; 2768 BP.

XX AC ABX71198;

XX AC ABX71198;

DT 05-MAR-2003 (first entry)

XX Novel human cDNA sequence #423.

Human; gene; ss; nervous system disorder; peripheral neuropathy;

Huntington's disease; amyotrophic lateral sclerosis; haemophilia;

neurodegenerative disease; Parkinson's disease; Alzheimer's disease;

autoimmune disease; systemic lupus erythematosus; rheumatoid arthritis;

insulin-dependent diabetes mellitus; anaemia; thrombocytopaenia; wound;

ulcer; burn; bone disorder; osteoporosis; osteoarthritis; stroke;

fibrosis; reperfusion injury; infection; allergic rhinitis; asthma;

coagulation disorder; cancer; tumour; inflammatory disease; septic shock;

Crohn's disease; anaphylaxis; proliferation; chemotactic;

differentiation; stem cell growth factor; haematopoiesis; chemokinetic;

haemostatic; antiinflammatory; expressed sequence tag; EST.

XX Homo sapiens.

XX WO200281731-A2.

XX 17-OCT-2002.

XX 29-JAN-2002; 2002WO-US001222.

XX 30-JAN-2001; 2001US-00774528.

XX (HYSE-) HYSEQ INC.

XX (GOOD/) GOODRICH R W.

XX Tang TY, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;

XX Xue AJ, Yang Y, Wehrman T, Wang J, Wang D, Drmanac RT;

XX WPI; 2003-058563/05.

XX Novel polypeptide useful for treating neurodegenerative diseases, myeloid

or lymphoid cell disorders, bone disorders, mechanical and traumatic

disorders, coagulation disorders, and inflammatory diseases.

XX Claim 1; Page; 612pp; English.

XX This invention relates to the cDNA sequences encoding an isolated novel

human polypeptide. The protein encoded by the nucleic acid of the

invention is useful for treating central and peripheral nervous system

diseases (e.g. peripheral neuropathy, Huntington's disease, amyotrophic

lateral sclerosis); neurodegenerative diseases (e.g. Parkinson's disease,

Alzheimer's disease); autoimmune disease (e.g. systemic lupus

erythematosus, rheumatoid arthritis, insulin-dependent diabetes mellitus)

; myeloid or lymphoid cell disorders (e.g. anaemia and thrombocytopoenia)
CC
CC ; wounds, ulcers, burns; bone disorders (e.g. osteoporosis,
CC osteoarthritis); mechanical and traumatic disorders (e.g. stroke, head
CC trauma); lung or liver fibrosis; reperfusion injury in various tissues;
CC bacterial, viral or fungal infections; allergic conditions such as
CC allergic rhinitis, asthma; coagulation disorders (e.g. haemophilia);
CC cancer and tumours; and inflammatory diseases (e.g. septic shock, Crohn's
CC disease, anaphylaxis). The protein may be used to inhibit the growth,
CC infection or function of infectious agents such as bacteria, fungi,
CC viruses, or to effect bodily characteristics, biorhythms or circadian
CC cycles of rhythms. The protein may also have
CC proliferation/differentiation, stem cell growth factor, hematopoiesis
CC regulation, immune stimulating or suppressing, chemotactic/chemokinetic,
CC haematostatic and thrombolytic, receptor/ligand, and antiinflammatory
CC activities. The cDNA sequences of the invention are useful for expressing
CC recombinant protein for analyses. The present sequence represents a novel
CC human cDNA sequence of the invention, this sequence is an expressed
CC sequence tag (EST) and was identified using subtractive hybridisation
XX
SQ Sequence 2768 BP: 456 A; 849 C; 995 G; 468 T: 0 U: 0 Other:

QY	786	GAGCTGTGAGGTGGCCAGGCCACAGACAGAGGTGACGTGGTGTACAGAGATGGGAGAGCT	845
Db	2591	GAGCTGTGAGGTGGCCAGGCCACAGATGGAGGTGACCTGGTACAGGACGGGAAGAGTT	2650
QY	846	GAGCTCCAGTTCSAAAGTCGCATAGAGGCTGCGGGCTGCATGCGGCAGCTGGTGGTGCA	905
Db	2651	GAGCTCCAGCTCSAAAGTGCACTGGAGGCTGTGGGCTGTATCGGAGGCTGGTGGTGCA	2710
QY	906	GCAGCGAGGCCAGGCAGATGCTGGGGAGTAGTACACCTGTGAGGCTGGGGG	953
Db	2711	GCAGGTGGCCAGGCAGACTCCGATAGTACAGCTGTGAAGCCAGGTG	2758
RESULT 4			
AD30509			
ID	AD30509	standard; cDNA; 1251 BP.	
XX	AC	AD30509;	
XX	DT	18-DEC-2003 (first entry)	
XX	DE	Human novel cDNA sequence, SEQ ID NO:591.	
XX	KW	Human; diagnostic; drug screening; forensics; gene mapping;	
XX	KW	biodiversity assessment; Parkinson's disease; Alzheimer's disease;	
XX	KW	neurodegenerative diseases; anaemia; platelet disorder; wound; burns;	
XX	KW	ulcers; osteoporosis; autoimmune disease; cancer;	
XX	KW	molecular weight marker; food supplement; antiparkinsonian; nootropic;	
XX	KW	neuroprotective; antianemic; anticoagulant; thrombolytic; vulnerary;	
XX	KW	antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;	
XX	KW	gene therapy; chromosome 17; gene; ss.	
OS	Hom	sapiens.	
XX	PN	WO2003029271-A2.	
XX	PD	10-APR-2003.	
XX	PF	24-SEP-2002; 2002WO-US030474.	
XX	PR	24-SEP-2001; 2001US-0324631P.	
XX	PA	(HYSE-) HYSEQ INC.	
XX	PI	Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;	
XX	PI	Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;	
XX	PI	Hailey-Vicente D, Drmanac RT;	
XX	PI	WPI; 2003-371981/35.	
XX	PI	P-FSDB; ADC31480.	
XX	PT	New polynucleotide and polypeptide useful for diagnosing, preventing or	
XX	PT	treating conditions such as neurodegenerative diseases, anemias, platelet	
XX	PT	disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or	
XX	PT	cancer.	
XX	PS	Claim 1; SEQ ID NO 591; 1185pp; English.	
XX	CC	The invention relates to 971 novel human cDNA sequences (ADC29919-	
XX	CC	ADC30898) and the polypeptides they encode (ADC30890-ADC31960). The	
XX	CC	invention also relates to nucleic acid sequences over 99% identical with	
XX	CC	the novel human cDNAs. The invention additionally encompasses expression	
XX	CC	vectors and host cells comprising a nucleic acid of the invention; the	
XX	CC	recombinant production of a polypeptide of the invention; an antibody	
XX	CC	against a polypeptide of the invention; a method of detecting	
XX	CC	polynucleotides or polypeptides of the invention; and methods of	
XX	CC	identifying a compound which binds to a polypeptide of the invention. The	
XX	CC	invention further discloses methods of preventing, treating or	
XX	CC	ameliorating a medical condition; kits comprising polynucleotide probes	
XX	CC	and/or monoclonal antibodies for carrying out the methods of the	
XX	CC	invention; methods for the identification of compounds that modulate the	
XX	CC	expression or activity of the polynucleotide and/or polypeptide; and 767	
XX	CC	contig sequences corresponding to the cDNA sequences of the invention	

CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: the sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIFO at
 CC ftp.wifo.int/pub/published_pct_sequences.

SQ Sequence 1251 BP; 329 A; 293 C; 363 G; 266 T; 0 U; 0 Other;

Query Match 21.8%; Score 218.6; DB 10; Length 1251;
 Best Local Similarity 86.1%; Pred. No. 8.2e-42;
 Matches 242; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 705 CACAGAGCCCAAGCAGTGTTCACAGAGCAGTGTGTCATATGAGTGGGACTGA 764
 DB 891 CCCAGAGCCCAAGTGTGTTCACAGAGCAGCAGCAGAGGAGGTGACAGCTGA 950
 QY 765 GCGAGGGGCCAGTGCACACTGAGCTGTGAGTGGCCAGGCCACAGACAGAGGTGACGTG 824
 DB 951 GCGAGGGGCCAGTGCACACTGAGCTGTGAGTGGCCAGGCCACAGACAGAGGTGACGTG 1010
 QY 825 GTACAAGGATGGGAAGAGCTGAGCTCCAGTTCGAAAGTGGCAGATAGAGGCTGGGGCTG 884
 DB 1011 GTACAAGGATGGGAAGAGCTGAGTTCAGCTCGAAAGTGGCGGTGGAGCGCTGGGGCTG 1070
 QY 885 CATCGGCGAGCTGGTGTGTCAGAGCGAGCCAGCCAGCAGATCTGGGAGTACACTGTGA 944
 DB 1071 CACAGGAGGTGGTGGTGTGTCAGAGCGGCGGCCAGCAGAGCGGGGAGTACAGCTGCGA 1130
 QY 945 GCGTGGGGCCAGCGGCTCTCCTCCAGCTGGATGTTTCAG 985
 DB 1131 GCGAGGGGGTACAGAGCTCTCCTCCGCTGCAGGTGGCAG 1171

RESULT 5
 AAH18453
 ID AAH18453 standard; cDNA; 2488 BP.
 XX AAH18453;
 AC AAH18453;
 XX
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human cDNA sequence SEQ ID NO:18549.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy; ss.
 KW
 XX Homo sapiens.
 OS
 XX EP1074617-A2.
 PN
 XX
 PD 07-FEB-2001.
 XX
 XX 28-JUL-2000; 2000EP-00116126.
 XX
 PR 29-JUL-1999; 93JP-00248036.
 PR 27-AUG-1999; 93JP-00300253.
 PR 11-JAN-2000; 2000JP-00118776.
 PR 02-MAY-2000; 2000JP-00183767.
 PR 09-JUN-2000; 2000JP-00241899.
 XX
 XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
 PT length cDNAs defined in the specification, and for the detection and/or
 PT diagnosis of the abnormality of the proteins encoded by the full-length
 PT cDNAs.
 XX
 PS Claim 8; SEQ ID NO 18549; 2537pp + Sequence Listing; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602 full-
 CC length cDNAs defined in the specification. Where a primer set comprises:
 CC (a) an oligo-dT primer and an oligonucleotide complementary to the
 CC complementary strand of a polynucleotide which comprises one of the 5602
 CC nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH24446 to AAH95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention
 XX
 SQ Sequence 2488 BP; 561 A; 637 C; 716 G; 574 T; 0 U; 0 Other;

Query Match 21.8%; Score 218.6; DB 4; Length 2488;
 Best Local Similarity 86.1%; Pred. No. 9.7e-42;
 Matches 242; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

QY 705 CACAGAGCCCAAGCAGTGTTCACAGAGCAGTGTGTCATATGAGTGGGACTGA 764
 DB 891 CCCAGAGCCCAAGTGTGTTCACAGAGCAGCAGCAGAGGAGGTGACAGCTGA 950
 QY 765 GCGAGGGGCCAGTGCACACTGAGCTGTGAGTGGCCAGGCCACAGACAGAGGTGACGTG 824
 DB 951 GCGAGGGGCCAGTGCACACTGAGCTGTGAGTGGCCAGGCCACAGACAGAGGTGACGTG 1010
 QY 825 GTACAAGGATGGGAAGAGCTGAGCTCCAGTTCGAAAGTGGCAGATAGAGGCTGGGGCTG 884
 DB 1011 GTACAAGGATGGGAAGAGCTGAGTTCAGCTCGAAAGTGGCGGTGGAGCGCTGGGGCTG 1070
 QY 885 CATCGGCGAGCTGGTGTGTCAGAGCGAGCCAGCCAGCAGATCTGGGAGTACACTGTGA 944
 DB 1071 CACAGGAGGTGGTGGTGTGTCAGAGCGGCGGCCAGCAGAGCGGGGAGTACAGCTGCGA 1130
 QY 945 GCGTGGGGCCAGCGGCTCTCCTCCAGCTGGATGTTTCAG 985
 DB 1131 GCGAGGGGGTACAGAGCTCTCCTCCGCTGCAGGTGGCAG 1171

RESULT 6
 ADR08183/c
 ID ADR08183 standard; cDNA; 3956 BP.
 XX ADR08183;
 AC ADR08183;
 XX
 DT 04-NOV-2004 (first entry)
 XX
 DE Full length human cDNA useful for treating neurological disease Seq 1689.
 KW gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;

osteoporosis; neurological disease; Alzheimer's disease;
 Parkinson's disease; dementia; short memory; cancer;
 sense or motor function; emotional reaction; fear response; panic;
 osteopathic; neuroprotective; nootropic; antiparkinsonian; cytostatic;
 tranquiliser.
 XX Homo sapiens.
 XX EP1447413-A2.
 XX 18-AUG-2004.
 XX 12-FEB-2004; 2004EP-00003145.
 XX 14-FEB-2003; 2003JP-00102207.
 XX 09-MAY-2003; 2003JP-00131452.
 XX (REAS-) RES ASSOC BIOTECHNOLOGY.
 XX Isogai T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T;
 PI Wakamatsu A, Ishii S, Nagai K, Irie R;
 PI WPI: 2004-583265/57.
 DR P-PSDB; ADR10139.
 XX New 1995 cDNA, useful for treating osteoporosis, neurological diseases,
 PT Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
 XX Claim 1; SEQ ID NO 1689; 2686pp; English.
 XX This invention relates to novel, isolated full length human cDNA
 CC molecules and the encoded proteins thereof. Specifically, it refers to
 CC cDNA clones obtained by an oligo-capping method, where none of these
 CC clones are identical to any known human mRNAs. The present invention
 CC describes an immunoassay to identify agonists and antagonists, as well as
 CC antibodies, antisense molecules and siRNAs that can all be used to bind
 CC to and modulate expression of the cDNA molecules. As such, these
 CC molecules are useful for diagnostic markers or therapeutic targets for
 CC the various diseases or morbid states. In particular, they are useful in
 CC gene therapy for treating osteoporosis, neurological disease, Alzheimer's
 CC disease, Parkinson's disease, dementia, short memory and various cancers,
 CC as well as for maintaining equilibrium of sense or motor function, and
 CC for treating emotional reaction, fear response and panic. Accordingly,
 CC they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,
 CC cytotatic and tranquiliser activities. This polynucleotide is a full
 CC length human cDNA sequence of the invention. NOTE: This sequence is not
 CC given in the sequence listing of the specification but can be obtained on
 CC CD-ROM from the European Patent Office, Vienna Sub-office.
 XX SQ Sequence 3956 BP; 882 A; 1211 C; 933 G; 930 T; 0 U; 0 Other;
 Query Match 21.8%; Score 218; DB 13; Length 3956;
 Best Local Similarity 84.5%; Pred. No. 1.5e-41;
 Matches 245; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
 QY 144 CCTGCATCATCAGAGCCCAAGATGATTTTGAAGAGAGCAGTCAGTCAATATGAGGT 203
 DB 1141 CTTGTCCATCCAGAGCCCAAGAGTGGTGTTCCTCAAGAGCAGCGGATGAGGT 1082
 QY 204 GCAGGCTCAGGCGGGGCCAGTGCCATGCTGAGCTGTGAGTGCAGCCAGCAGCGGA 263
 DB 1081 CGAGCGGAGCGGGGGCCAGTGCCATGCTGAGCTGTGAGTGCAGCCAGCAGATGGA 1022
 QY 264 GGTGACGTGGTACAGGATGGGAAGAGCTGAGCTCCAGCTCAAAAGTGGGCATGGAGGT 323
 DB 1021 GGTGACATGGTACAGGAGCGGAAGAGCTGAGCTCCAGCTCAAAAGTGGGCATGGAGG 962
 QY 324 CAAGAGGTGCACAGCAGGCTGGTGTCCACAGCGCGGCAAGCAGATGCTGGGGAGTA 383
 DB 961 CAGCGGCTACACAGGAGGCTGGTGTGACAGCAGCGCGGCGAGTGTCTGGGGAGTA 902
 QY 384 CAGCTGTGAGCTGGGGGCCAGAGAGTCTCTCCACTGCACATCAG 433

Db 901 CAGCTGTGAGCGGGGGGCCAACAGCTGTCTCTCCGCTGCACGTGGCAG 852
 RESULT 7
 AAI59251
 ID AAI59251 standard; cDNA; 2737 BP.
 XX AC AAI59251;
 XX 22-OCT-2001 (first entry)
 XX Human polynucleotide SEQ ID NO 1454.
 XX Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
 KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia; ss.
 KW Homo sapiens.
 XX WO200153312-A1.
 XX 26-JUL-2001.
 XX 26-DEC-2000; 2000WO-US034263.
 XX 23-DEC-1999; 99US-00471275.
 PR 21-JAN-2000; 2000US-00488725.
 PR 25-APR-2000; 2000US-00552317.
 PR 20-JUN-2000; 2000US-00598042.
 PR 19-JUL-2000; 2000US-00620312.
 PR 03-AUG-2000; 2000US-00653450.
 PR 14-SEP-2000; 2000US-00682191.
 PR 19-OCT-2000; 2000US-00693036.
 PR 29-NOV-2000; 2000US-00727344.
 XX (HVSE-) HVSEQ INC.
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
 PI Zhou P, Goodrich R, Drmanac RT;
 XX WPI: 2001-442253/47.
 DR P-PSDB; AAM40095.
 XX Novel nucleic acids and polypeptides, useful for treating disorders such
 PT as central nervous system injuries.
 XX Claim 1; SEQ ID NO 1454; 10078pp; English.
 XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
 CC encoded polypeptides (AAM38642-AAM42213) with nootropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders. Note: The sequence data for this patent did not form
 CC part of the printed specification
 XX SQ Sequence 2737 BP; 484 A; 840 C; 899 G; 514 T; 0 U; 0 Other;
 Query Match 14.0%; Score 140.2; DB 4; Length 2737;
 Best Local Similarity 48.4%; Pred. No. 3.3e-23;
 Matches 484; Conservative 0; Mismatches 508; Indels 9; Gaps 3;

Db 1996 GAGGTGGACCGAGAGGACGCCCTCTGTGCGTTGGTAAAGGACGGGAGAGGTGGAGGAG 2055
QY 301 AGCTCAAAAGTGGCATGAGAGTCAAAAGGTTGCACACGAGGCTGGTGTGCGACACAGCGG 360
Db 2056 AGTGACTTGTGCTGGAGATGAGGGGCCCATCGCCGCTGGTGTGCTGCGCCGAC 2115
QY 361 GGAAGACAGATGCTGGGAGTACAGCTGTGTAGGTGGGGGCCAGAGAGTCTCTTTCAC 420
Db 2116 CAGCCCTCAGACGGGGGCGAGTTTCAGTGTGCTGCGTGGAGATGAGTGTGCTTACT 2175
QY 421 CTGCACATCAGACGCCCAAGGGGTGTTTGGAAAGGACGACTAGTCATTAATGAGTG 480
Db 2176 GTACCATCAGACAGCTCTCTCGTGGATCG---TGTATCCAGCGGCAAGGTGTATGTG 2232
QY 481 CAGGCTGAGGCGGGACCACTGCCATGCTGAGCTGTGAGTGGCCGACGCCCCAGACGAG 540
Db 2233 CGAGCGTGGCGCTGGAGCGTGTGCTGACCTGTGAGCTATGCGGCCCTGGGCAGAG 2292
QY 541 GTGACGTGTACAGGACGGGAAGAGTGTGAGTCCAGCTCAAAAGTACGATGGAGTGC 600
Db 2293 GTGCGTGTGACCAAGGATGGAGAGGTGGTGAGAGCCCGCGCTGCTCTGCGAAG 2352
QY 601 AAGGCTGCACACGAAGGCTGTGCTGACAGAGTGGGCAAGCAGATGCTGGGAGTAC 660
Db 2353 GAAGACACTGTCCCGCGCTGTGCTGCGCGTGTCCAGCTCGAGACTCCGGCGAGTAC 2412
QY 661 AGCTCGAGGCTGGGGCCAGAGAGTCTCTTTCAACTGCAATCAGACAGGCCCAAGGCA 720
Db 2413 TTGTGTGAATGTGACATGAGTGGCGCTCTTCACTGTACCGTACAGAACCC---CCA 2459
QY 721 GTGTTTGCAGGAGCAGTGTGTGCATTAATGAGTGTGCGACTGAGGCGAGGGCCAGTGCC 780
Db 2470 GTGCGGATCATATACCTCGCATGAGGTGACCTTGTATCGCGTGTGAGTGTGTG 2529
QY 781 ACACGTGAGTGTGAGGTGGCCAGGCCAGACAGAGTGTGCTGTACAGATGGGAG 840
Db 2530 GTGCTGATGTGTAATCTGTCTCGGAGGATGCCCCGTGTGCGTGTGTAACAAGATGGGCTG 2589
QY 841 AAGCTGAGTCCAGTTCGAAAGTGGCATAGAGGCTGGGGTGTGATCGCGCAGCTGTGTG 900
Db 2590 GAAGTGGAGGAGAGCGAGGCGCTGTGTGTGAGAGGATGGCCACGCTGCGCTGTGTG 2649
QY 901 GTGACGAGGAGCGCCAGCGAGATCTCTGGGAGTACACTGTGAGGCTGGGGGCCAGCGG 960
Db 2650 CTACCTGTGTCTAGCCCGAGGACGGGGCGAGTTTGTATGTGTGCTGTGAGATGACTCG 2709
QY 961 CTCTCTTCCACCTGATGTTTTCAGAGCCCAAGCGGTGTT 1001
Db 2710 GCCTTCTTCACTGTCACTGTACAGAGCCCTCCAGTGCAGTT 2750

RESULT 9
ID ADD14722 standard; cDNA; 5382 BP.
XX AC ADD14722;
XX 01-JAN-2004 (first entry)
XX Human src biomarker polynucleotide SEQ ID NO:116.
XX predictor set; protein tyrosine kinase activity modulator;
KW protein tyrosine kinase pathway; protein tyrosine kinase; cytostatic;
KW gene therapy; drug sensitivity; genetic profile; cancer; human; gene; ss.
XX Homo sapiens.
XX WO2003062395-A2.
XX 31-JUL-2003.
XX 17-JAN-2003; 2003WO-US001981.

PR 18-JAN-2002; 2002US-0350061P.
XX (BRIM) BRISTOL-MYERS SQUIBB CO.
XX Huang F, Fairchild CR, Lee FY, Shaw P;
XX WPI; 2003-636735/60.
DR P-PSDB; ADD14123.
XX New polynucleotides and polypeptides for predicting the activity of
PT compounds that interact with protein tyrosine kinases and/or protein
PT tyrosine kinase pathways.
XX Claim 2; SEQ ID NO 116; 139pp; English.
XX The present invention describes a predictor set comprising a plurality of
XX polynucleotides or polypeptides whose expression pattern is predictive of
CC the response of cells to treatment with a compound that modulates protein
CC tyrosine kinase activity or members of the protein tyrosine kinase
CC pathway. Also described: (1) predicting whether a compound is capable of
CC modulating the activity of cells, comprising obtaining a sample of cells,
CC determining whether the cells express a plurality of markers, and
CC correlating the expression of the markers to the compound's ability to
CC modulate the activity of the cells; (2) a plurality of cell lines for
CC identifying polynucleotides and polypeptides whose expression levels
CC correlate with compound sensitivity or resistance of cells associated
CC with a disease state; and (3) identifying polynucleotides and
CC polypeptides that predict compound sensitivity or resistance of cells
CC associated with a disease state, comprising subjecting the plurality of
CC cell lines to one or more compounds, analysing the expression pattern of
CC a microarray of polynucleotides or polypeptides, and selecting
CC polynucleotides or polypeptides that predict the sensitivity or
CC resistance of cells associated with a disease state by using the
CC expression pattern of the microarray. The polynucleotides and
CC polypeptides have cytostatic activities, and can be used in gene therapy.
CC The polynucleotides and polypeptides are useful in predicting the
CC activity of compounds that interact with protein tyrosine kinases and/or
CC protein tyrosine kinase pathways. These may be used in determining drug
CC sensitivity in patients to allow the development of individualized
CC genetic profiles which aid in treating diseases and disorders (e.g.
CC cancer) based on patient response at a molecular level. The present
CC sequence is used in the exemplification of the present invention.
XX SQ Sequence 5382 BP; 943 A; 1699 C; 1798 G; 942 T; 0 U; 0 Other;
Query Match 14.0%; Score 140.2; DB 10; Length 5382;
Best Local Similarity 48.4%; Pred. No. 3.9e-23;
Matches 484; Conservative 0; Mismatches 508; Indels 9; Gaps 3;
QY 1 AAAGATGGGAAGAGCTGAGCTCCAGCTTGAAAGTGCATGTAGAGGCAAGGCTGCAGA 60
Db 1760 AAGGATGGGCAGAAGGTGGAGGAGAGCGAGTTGCTGTGTGAAGATGGATGGCGCAA 1819
QY 61 CGGAGGCTGTGTGTCAGAGGAGGAGCAAGCGGATGCCGGGACATACAGTGGAGGCC 120
Db 1820 CACCGTCTGATCTCCTGCTGAGGCGCAAGTCCAGACAGTGGCGGTTTGTAGTGCAGACA 1879
QY 121 AGGGGCCAGAGGGTCTCTTCCGCTGCACATCACAGAGCCCAAGATGATGTTTGCAAAG 180
Db 1880 GAAGGGGTCTCGGCCCTTCTTCGGCGTCACTGTCAAGATCTCCGTCGCACATCG---TG 1936
QY 181 GAGCAGTCACTGATTAATGAGTGCAGGCTGAGCGGGGCCAGTGCCTGCTGAGCTGT 240
Db 1937 GACCCCGAGAACATGTGTTCGTGTCATGCCATAACTCCAGTGTGTGTCATGCTGGCCTGT 1996
QY 241 GAGTGGCCAGGCCCCAGACGAGGTGACGTGTGTCACAGATGGGAGAGCTGAGTCC 300
Db 1997 GAGGTGACCGAGAGGACGCCCTCTGTGCTGTGTCACAGGACGGGAGGAGGAG 2056
QY 301 AGCTCAAAAGTGGGCATGGAGGTCAAAAGGTTGCACACGAGGCTGGTGTGCTCACAGCGG 360
Db 2057 AGTGACTTGTGCTGGAGATGAGGGGCCCATCGCCGCTGTGCTGCTGCCGCCACC 2116

Db 2866 TGGTCTGATGTGAATCTCTCGGAGGATGCCCTGTGCGCTGGTACAAAGATGGC 2807
Qy 839 AGAAGCTGAGCTCCAGTTTCGAAGTGGCGATAGAGGCTGCGGGCTGTCATGCGGACAGTGG 898
Db 2806 TGGAGTGGAGGAGAGCGAGGCCCTGGTCTGCGAGAGGATGGGCCACGCTGCCCGCTGG 2747
Qy 899 TGGTGCAGCAGCAGCGCCAGGCGAGATGCTGGGGAGTACACCTGTGAGGCTGGGGCCAGC 958
Db 2746 TGCTACCTTGCTCAGCCGAGGACGGGGGCGAGTTTGTATGTATGATGCTGAGATGACT 2687
Qy 959 GGCTCTCTCCACCTGGATGTTTCAGAGCCCAAGCGGTGTT 1001
Db 2686 CGGCTTCTTCACTGCTCAGCTGTCACAGAGCCCTCCAGTGCAGTT 2644

RESULT 12
AAF24162
ID AAF24162 standard; DNA; 1645 BP.
XX
AC AAF24162;
XX
DT 23-MAR-2001 (first entry)
XX
DE Human secreted protein DNA #12.
XX
KW Secreted protein; gene therapy; vaccine; cancer; leukemia;
KW autoimmune disease; allergy; inflammation; graft rejection;
KW hyperproliferation; cardiovascular; infection; ds.
XX
OS Homo sapiens.
XX
PN WO20007375-A1.
XX
PD 14-DEC-2000.
XX
PF 02-JUN-2000; 2000WO-US015187.
XX
PR 07-JUN-1999; 99US-0137725P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Birse CE, Duan RD, Soppet DR, Rosen CA, Shi Y;
PI Lafleur DW, Olsen HS, Ebner R, Florence KA, Ni J, Young PE;
XX WPI; 2001-061741/07.
XX
XX Nucleic acids encoding 26 human secreted polypeptides, useful e.g. for
PT preventing, diagnosing and/or treating cancers and for promoting wound
PT healing.
XX
PS Claim 1; Page 425-426; 530pp; English.
XX
CC The present invention relates to 26 secreted human proteins. The proteins
CC may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. For example, they
CC may be used in gene therapy or in vaccines. Typical of diseases which are
CC potentially treatable are cancers (including leukemia), autoimmune
CC diseases, allergies, inflammation, graft rejection, hyperproliferation,
CC cardiovascular diseases (particularly critical limb ischemia and coronary
CC disease) and any involving abnormal angiogenesis, neurodegeneration
CC and/or infectious diseases
XX
SQ Sequence 1645 BP; 338 A; 474 C; 506 G; 327 T; 0 U; 0 Other;

Query Match 12.3%; Score 123.2; DB 5; Length 1645;
Best Local Similarity 49.5%; Pred. No. 3.1e-19;
Matches 376; Conservative 0; Mismatches 378; Indels 6; Gaps 2;

Qy 230 TGCTGAGCTGAGTGCCCGCCAGCGAGGTGAGTGGTACAAAGATGGGAAGA 289
Db 591 TGCTGAGCTGTGAGCTCTCAAGGTTGACTTCCCGGCAACCTGGTACAAAGATGGGAGA 650
Qy 290 AGCTGAGCTCCAGCTCAAAAGTGGCATGGAGGTCAAGGGTGCACCGGAGGTGGTGC 349

Db 651 AGGTGGAGGAGCGAGTTGCTGGTGGTGAAGATGGATGGGCGCAACACCGCTCTGATCC 710
Qy 350 TGCCACAGCGCGGCAAAAGCAGATGCTGGGGAGTACAGCTGTGAGGCTGGGGCCAGAGAG 409
Db 711 TGCTGAGGCGCAAAAGTCCAGGACAGTGGCGAGTTTGAAGTGCAGGACAGAGGGGTCTCG 770
Qy 410 TCTCTTCCACTGCACATACAGAGCCCAAGGGGTGTTTTCGGAAGAGAGTCAAGTGC 469
Db 771 CTTTCTTCGGGCTCACTGTCCAAGATCTCCCGTGCACATCG---TGGACCCCGAGAAC 827
Qy 470 ATAATGAGGTGAGGCTGAGGCGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCAGC 529
Db 828 ATGTGTTCTGTCATGCCATAACTTCCGAGTGTGTCATGCTGCCCTGTGAGGTGGACCGAG 887
Qy 530 CCCAGACGGAGGTGACGTTGTTACAAAGGACGGGAAGAGCTGAGCTCCAGCTCAAAAGTAC 589
Db 888 AGGACGCCCTGTGCTTGGTGAAGGACGGGAGGAGTGGAGGAGTGAATTCGTTGG 947
Qy 590 GCATGGAGGTCAAGGGCTGCACAGAGGCTGGTGTGTCAGCAGAGTGGGCAAGAGATG 649
Db 948 TGCTGGAGAAATGAGGGGGCCCATCGCGCTGTGCTGCGCAGCGCTCCAGACG 1007
Qy 650 CTGGGAGTACAGCTGCGAGGCTGGGGCCAGAGAGTCTCTTTCAACTGCACATCACAG 709
Db 1008 GGGGAGGTTTCAAGTGCCTGCGATGAGTGAAGTGTGCTTCACTGTCACTATCACAG 1067
Qy 710 AGCCCAAGGACAGTGTGTTGCCAAGGAGCAGTTGGTGCATAAATGAGTGGCGACTGAGGCAG 769
Db 1068 A---CGTCTCTCTGTTGATCGTATCCAGCGGCAAGGTGTATGTGGCAGCGTGGCC 1124
Qy 770 GGGCAGTGCACACTGAGCTGTGAGTGGGCCAGAGGCTCCAGAGTGGTGTGAGTGTGAGTACA 829
Db 1125 TGGAGCGTGTGCTGCTGACCTGTGAGCTATGCCGGCCCTGGGCGAGGTGGCTGGACCA 1184
Qy 830 AGGATGGGAAGAGCTGAGCTCCAGTTCGAAGTGGCATAGAGGCTCGGGCTCATGCG 889
Db 1185 AGGATGGAGAGGAGTGTGGAGAGCCCGCGCTGCTCTCGAAGGAAGACATGTGTC 1244
Qy 890 GGCAGCTGGTGTGCAGCAGCGCCAGGCGAGATGCTGGGGAGTACACCTGTGAGGCTG 949
Db 1245 GCCGCTGGTGTGCCCGCTGTCCAGTCCAGGACTCCGGCGAGTACTTGTGTGAATTG 1304
Qy 950 GGGGCGAGCGGCTCTCTTCCACCTGGATGTTTCAGAGCC 989
Db 1305 ACGATGAGTGGCGCTCTCTTCACTGTCAACCGTCACAGAGTC 1344

RESULT 13
ABV74346
ID ABV74346 standard; cDNA; 1948 BP.
XX
AC ABV74346;
XX
DT 21-FEB-2003 (first entry)
XX
DE Human IL-1RRP3 encoding cDNA SEQ ID NO 1.
XX
KW Human; interleukin 1 receptor related protein 3; IL-1RRP3; inflammation;
KW arthritis; septicaemia; autoimmune disease; transplant rejection;
KW infection; stroke; AIDS; osteoporosis; cancer; Alzheimer's disease;
KW anti-inflammatory; antiarthritic; antibacterial; immunosuppressive;
KW cerebroprotective; vasotropic; antiviral; osteopathic; cytostatic;
KW antiarteriosclerotic; nootropic; neuroprotective; chromosome 2q33; gene;
ss.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT 413..1699
FT /*tag= a
FT /product= "IL-1RRP3"
XX

PN WO9918990-A1.
 XX PD 22-APR-1999.
 XX PF 09-OCT-1998; 98WO-US021347.
 XX PR 10-OCT-1997; 97US-0061720P.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX PI Young PR;
 XX DR WPI; 1999-277442/23.
 XX DR P-PSDB; ABB98343.
 XX PT Novel interleukin 1 receptor related protein 3 (IL-1RRP3) or variants
 XX PT for, e.g. treatment of Alzheimer's disease.
 XX PS Claim 2; Page 30-31; 43pp; English.
 XX CC The invention relates to polynucleotides encoding interleukin 1 receptor
 CC related protein 3 (IL-1RRP3) or variants. The polynucleotides are used to
 CC diagnose a disease or susceptibility to a disease in a subject related to
 CC expression or activity of the IL-1RRP3 to screen for compounds which
 CC stimulate or inhibit the function of IL-1RRP3 or to treat a subject in
 CC need of enhanced or inhibited IL1RRP3 activity. Specifically mentioned in
 CC the specification is the use of the IL-3RRP3 polynucleotide or
 CC polypeptide for treatment of chronic and acute inflammation, arthritis,
 CC septicemia, autoimmune disease, transplant rejection, graft-versus-host
 CC disease, infection, stroke, ischemia, acute respiratory disease
 CC syndrome, restenosis, brain injury, AIDS, bone diseases e.g.
 CC osteoporosis, cancer, atherosclerosis and Alzheimer's disease. The
 CC present sequence is that of a human IL-1RRP3 encoding cDNA of the
 CC invention. The IL-1RRP3 gene is located on chromosome 2q33
 XX
 SQ Sequence 1948 BP; 397 A; 561 C; 600 G; 386 T; 0 U; 4 Other;
 Query Match 12.3%; Score 123.2; DB 2; Length 1948;
 Best Local Similarity 49.5%; Pred. No. 3.2e-19;
 Matches 376; Conservative 0; Mismatches 378; Indels 6; Gaps 2;

QY 230 TGCTGAGCTGTGAGTGGCCAGCCAGCGAGGTGACGTGTACAGGATGGGAGA 289
 DB 825 TGCTGACTGTGAGCTCTCAAGGGTGGACTTCCCGGCAACCTGGGTACAGGATGGCAGA 884
 QY 290 AGCTGAGCTCCAGCTCAAAAGTGGSCATGGAGGTCAAAGGTGCACACGGAGGCTGTGTC 349
 DB 885 AGTGGAGGAGCGAGTGTCTGTGTGTAAGATGGATGGGCGCAAAACCCGTCTGATCC 944
 QY 350 TGCCACAGCGCGGCAAGCAGATGCTGGGGAGTACAGCTGTGAGGCTGGGGCCAGAGAG 409
 DB 945 TGCTGAGGCCAAAGTCCAGGACAGTGGGAGTTGATGTCAGGACAGAGGGTCTCGG 1004
 QY 410 TCTCCTTCCACTGCACATCAGAGCCCAAGGGGTGTTTGGGAAGAGCAGTCAGTGC 469
 DB 1005 CTTTCTTCGGGGTCACTGTCCAAAGATCTCCCGTGCACATCG---TGGACCCCGAGAAC 1061
 QY 470 ATAATGAGGTGTCAGGCTGAGCGGGGACCACTGCCATGCTGAGGTGTGAGGTGGCCAGC 529
 DB 1062 ATGTGTTGTCATGCCATTAATCTCCGAGTGTGTATGCTGCTGCTGTGAGGTGACCGAG 1121
 QY 530 CCCAGACGGAGGTGACGTGGTGTACAGGACGGGAAGAAGCTGAGCTCCAGCTCAAAAGTAC 589
 DB 1122 AGGACGCCCTGTGCGTTGGTTGATACAGGACGGGAGGAGGTGAGGAGTCACTTCGTGG 1181
 QY 590 GCATGGAGTCAAGGCTGCACAGAGGCTGGTAGTGCACAGAGTGGGCAAGCAGATG 649
 DB 1182 TGCTGGGAATGAGGGGGCCCCATCGCCGCTGGTGTGTCGCCGCCACCCAGCCCTCAGACG 1241
 QY 650 CTGGGGAGTACAGCTGCGAGGCTGGGGGCCAGAGAGTCTCTTTCAACTGCACATCACAG 709
 DB 1242 GGGCGGAGTTTCAGTGCCTCGCTGGAGATGAGTGTGCTTACTTCACTGTCACTACAG 1301

QY 710 AGCCCAAGGCGAGTGTGTTGCCAAGGAGCAGTGTGTCATAAATGAGGTGCGGACTGAGGCAG 769
 DB 1302 A---CGTCTCCTCGTGTGATCGTGTATCCAGCGGCAAGGTGTATGTGGCAGCCGTCGCC 1358
 QY 770 GGCCAGTGCACACTGAGCTGTGAGGTGGCCAGGCCAGAGGTGACGTGGTACA 829
 DB 1359 TGGAGCGTGTGCTGCTGACCTGTGAGCTATGCGCGCCCTGGGCGAGGTGCGCTGGACCA 1418
 QY 830 AGGATGGGAAGAAGCTGAGCTCCAGTTCGAAAGTGGCGATAGAGGCTGCGGGCTGCATGC 889
 DB 1419 AGGATGGAGAGGAGGTGGTGAGAGCCCGCGCTGCTCTGCACAAGAGACACTGTCC 1478
 QY 890 GGCAGCTGGTGTGCAGCAGCAGCCAGCAGATGCTGGGGAGTACACCTGTGAGGCTG 949
 DB 1479 GCCGCTGGTGTGCTGCGCGCTGTCCAGCTCGAGGACTCCGGCGAGTACTTGTGTGAAATTG 1538
 QY 950 GGGGCCAGCGGCTCTCTTCCACCTGGATGTTTCAGAGCC 989
 DB 1539 ACGATGAGTGGCGCTCTCTTCACTGTCCACCGCTCACAGAGTC 1578

RESULT 14
 ABV74347
 ID ABV74347 standard; cDNA; 1949 BP.
 XX AC ABV74347;
 XX DT 21-FEB-2003 (first entry)
 XX DE Human IL-1RRP3 encoding cDNA SEQ ID NO 3.
 XX KW Human; interleukin 1 receptor related protein 3; IL-1RRP3; inflammation;
 KW arthritis; septicemia; autoimmune disease; transplant rejection;
 KW infection; stroke; AIDS; osteoporosis; cancer; Alzheimer's disease;
 KW anti-inflammatory; antiarthritic; antibacterial; immunosuppressive;
 KW cerebroprotective; vasotropic; antiviral; osteopathic; cytostatic;
 KW antiarteriosclerotic; nootropic; neuroprotective; chromosome 2q33; gene;
 KW ss.
 XX OS Homo sapiens.
 XX FH Key Location/Qualifiers
 FT CDS 414..1700
 FT /*tag= a
 FT /product= "IL-1RRP3"
 XX PN WO9918990-A1.
 XX PD 22-APR-1999.
 XX PF 09-OCT-1998; 98WO-US021347.
 XX PR 10-OCT-1997; 97US-0061720P.
 XX PA (SMIK) SMITHKLINE BEECHAM CORP.
 XX PI Young PR;
 XX DR WPI; 1999-277442/23.
 XX DR P-PSDB; ABB98344.
 XX PT Novel interleukin 1 receptor related protein 3 (IL-1RRP3) or variants
 XX PT for, e.g. treatment of Alzheimer's disease.
 XX PS Claim 13; Page 31-32; 43pp; English.
 XX CC The invention relates to polynucleotides encoding interleukin 1 receptor
 CC related protein 3 (IL-1RRP3) or variants. The polynucleotides are used to
 CC diagnose a disease or susceptibility to a disease in a subject related to
 CC expression or activity of the IL-1RRP3 to screen for compounds which
 CC stimulate or inhibit the function of IL-1RRP3 or to treat a subject in
 CC need of enhanced or inhibited IL1RRP3 activity. Specifically mentioned in
 CC the specification is the use of the IL-3RRP3 polynucleotide or

Matches 376; Conservative 0; Mismatches 378; Indels 6; Gaps 2;	
QY 230	TGCTGAGCTGTGAGTGGCCAGGCCACAGACGGAGGTGACGTGGTACAAAGGATGGGAAGA 289
Db	
1082	TGCTGACTTGTGAGCTCTCAAGGTGGACTTCCCGGCAACCTGGTACAAAGGATGGGCAGA 1141
QY	290 AGCTGAGCTCCAGCTCAAAAGTGGGCATGGAGGTCAAAGGGTGCACACGGAGGCTGGTGCC 349
Db	
1142	AGGTGGAGGAGAGCGAGTGTGCTGGTGAAGATGGATGGGCGCAACACACCGTCTGTGATCC 1201
QY	350 TGCCACAGCGGGCCAAAGACAGATGCTGGGGAGTACAGCTGTGAGCTGGGGCCACAGAGAG 409
Db	
1202	TGCTTGAGGCCAAAGTCCAGGACAGTGGCGAGTTTGAAGTGACAGGACAAAGGGGTCTCGG 1261
QY	410 TCTCCTTCCACCTGCACATCACAGAGCCCAAGGGGTGTTTGCGAAGGAGCAGTCAGTGCC 469
Db	
1262	CCTTCTTCGGCGTCACTGTCCAAGATCTCCCGTGCACATCG---TGGACCCCGAGAAC 1318
QY	470 ATAATGAGGTGAGGCTGAGCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCAGC 529
Db	
1319	ATGTGTCGTGCATGCCATACTTCCGAGTGTGTCACTGGCCTGTGAGGTGGACCCGAG 1378
QY	530 CCCAGACGGAGTACGCTGTACAGGACGGGAAGCTGAGCTCCAGCTCAAAAGTAC 589
Db	
1379	AGGACGCCCTGTGCTGTGTTGTTACAAAGACGGCGAGAGGTGGAGGAGTGACTTCGTGG 1438
QY	590 GCATGGAGTCAAGGGCTGCACACGAAGGCTGGTAGTGCAGCAGGTGGGCAAAAGCAGATG 649
Db	
1439	TGCTGGAGAATGAGGGGCCCCATCGCCGCTGGTGTCTGCCGCCACCCAGGCCCTCAGACG 1498
QY	650 CTGGGGAGTACAGCTGCGAGGCTGGGGGCCAGAGAGTCTCCTTTCAACTGCACATCAACAG 709
Db	
1499	GGGGCGAGTTTCAGTGCCTGCTGGAGATGAGTGTGCCCTACTTCACTGTCAACCATCACAG 1558
QY	710 AGCCCAAGGCAGTGTTCGCAAGGACAGTGGTGTGCATAATGAGGTGCGGACTGAGGCAG 769
Db	
1559	A---CGTCTCTCTGGGATCGTGATATCCAGCGGCAAGGTGTATGTGGCAGCCGTGGGCC 1615
QY	770 GGGCCAGTGCCACACTGAGCTGTGAGGTGGGCCACAGGCCCCAGACAGAGGTGACGTGGTACA 829
Db	
1616	TGGAGCGTGTGGTGTGACCTGTGAGCTATGCCGCCCTGGGCAGAGGTGCGCTGGACCA 1675
QY	830 AGGATGGGAAGAAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGGCTCGGGGCTGCATGC 889
Db	
1676	AGGATGGAGAGAGGAGTGGTGGAGAGCCCGCGCTGCTCTCTGCAGAAAGAGACACTGTGCC 1735
QY	890 GGCAGCTGGTGTGCAGCAGCAGCGCCAGGCAGATGCTGGGGAGTACACCTGTGAGGCTG 949
Db	
1736	GCCGCTGGTGTGCTGCCCGCTGTCCAGCTCGAGGACTCCGGCGAGTACTTTGTGAATTG 1795
QY	950 GGGGCCACGGCTCTCTTCCACCTGGATGTTTCAGAGCC 989
Db	
1796	ACGATGAGTCGGCCCTCCTTCACTGTCAACCGTCAACAGATC 1835

Search completed: March 20, 2005, 18:54:48
Job time : 588.044 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 13:15:10 ; Search time 581.044 Seconds
(without alignments)
10198.307 Million cell updates/sec

Title: US-10-077-130-4_COPY_10500_11500

Perfect score: 1001

Sequence: 1 gcctctgtgagtgaggaa.....ttcctagaagatgtgaaaaa 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_16Dec04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	20489	12 ADQ22881	Adq22881 Human sof
2	1001	100.0	24120	8 ABX11642	Abx11642 Human ser
3	997.8	99.7	7564	8 ABZ24581	Abz24581 Human cel
4	997.8	99.7	14061	6 ABV99363	Abv99363 Human NOV
5	997.8	99.7	14109	6 ABV99362	Abv99362 Human NOV
6	343.4	34.3	352	5 AAF64470	Aaf64470 Novel hum
7	265	26.5	642	6 ABQ61169	Abq61169 Obscurin
8	265	26.5	707	6 ABQ61180	Abq61180 Obscurin
9	238.8	23.9	2155	6 ABK99965	Abk99965 DNA encod
10	200.6	20.0	898	6 ABL58139	AbL58139 Human pho
11	107.4	10.7	2534	11 ADM03261	Adm03261 Human CDN
12	86	8.6	1005	6 ABN21414	Abn21414 Human ORF
13	62	6.2	3935	11 ADM01849	Adm01849 Human CDN
14	60	6.0	60	6 ABN40768	Abn40768 Human spl
15	59	5.9	93801	9 ABX13540	Abx13540 Human RGS
16	59	5.9	103052	13 ADQ99963	Adq99963 Antagonis
17	58.8	5.9	3999	4 AAI61037	Aai61037 Human pol
18	57.8	5.8	2737	4 AAI59251	Aai59251 Human pol
19	57.8	5.8	5378	13 ADR67197	Adr67197 Human bla
20	57.4	5.7	31595	10 ADF81661	Adf81661 Leukaemia

21	57.4	5.7	81940	4 AAS05390	Aas05390 Human tit
22	57.4	5.7	81940	6 ABK64829	Abk64829 Human ben
23	57.4	5.7	81940	12 ADQ17315	Adq17315 Human sof
24	56.4	5.6	5382	10 ADD14722	Add14722 Human src
25	53.6	5.4	1948	2 ABV74346	Abv74346 Human IL-
26	53.6	5.4	1949	2 ABV74347	Abv74347 Human IL-
27	53.6	5.4	2170	6 ABQ54970	Abq54970 Human ova
28	52	5.2	1645	5 AAF24162	Aaf24162 Human sec
29	50.8	5.1	1337	2 AAZ17263	Aaz17263 Human gen
30	50.6	5.1	1218	3 AAA02488	Aaa02488 Human col
31	50.4	5.0	2000	8 ADA71938	Ada71938 Rice gene
32	49.8	5.0	2254	3 AAC93403	Aac93403 Human sec
33	49.8	5.0	2254	5 AAF24183	Aaf24183 Human sec
34	48.6	4.9	10732	3 AAA10594	Aaa10594 Gene enco
35	47.6	4.8	716	6 ABK35726	Abk35726 CDNA sequ
36	44	4.4	1435	8 ABV74143	Abv74143 Human dua
37	43.8	4.4	1000	3 AAA02484	Aaa02484 Human col
38	43.6	4.4	5452	10 ADC86736	Adc86736 Human GPC
39	43	4.3	7328	4 AAK89081	Aak89081 Human dig
40	43	4.3	20565	4 AAK89083	Aak89083 Human dig
41	43	4.3	20565	4 AAS28908	Aas28908 Human inm
42	43	4.3	20565	4 AAL03357	Aal03357 Human rep
43	43	4.3	20565	10 ADB31749	Adb31749 Human nov
44	42.4	4.2	1593	3 AAA02504	Aaa02504 Human col
45	42.4	4.2	1945	13 ADR96750	Adr96750 Human dua

ALIGNMENTS

RESULT 1
ADQ22881
ID ADQ22881 standard; DNA; 20489 BP.
XX
AC ADQ22881;
XX
DT 26-AUG-2004 (first entry)
XX
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5701.
XX
KW soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human; ds.
XX
OS Homo sapiens.
XX
PN WO2004048938-A2.
XX
PD 10-JUN-2004.
XX
PF 26-NOV-2003; 2003WO-US038193.
XX
PR 26-NOV-2002; 2002US-0429739P.
XX
(PROT-) PROTEIN DESIGN LABS INC.

PI Aziz N, Ginsburg WM, Zlotnik A;
XX
DR WPI; 2004-441208/41.
XX
PT Early detection of soft tissue sarcoma comprises determining expression
of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
sarcoma.
XX

PS Example 2; SEQ ID NO 5701; 210pp; English.

XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
presence of soft tissue sarcoma. The method of the invention has

CC	cytostatic applications and may be useful for detecting soft tissue									
CC	sarcoma, possibly via gene therapy or vaccine production. The nucleic									
CC	acid sequences may be useful in diagnostic and screening applications.									
CC	The current sequence is that of a human soft tissue sarcoma-upregulated									
CC	DNA of the invention. The current sequence is not shown within the									
CC	specification per se but was submitted in CD format by the inventor.									
XX										
SQ	Sequence 20489 BP; 3917 A; 5815 C; 6983 G; 3225 T; 0 U; 549 Other;									
	Query Match	100.0%;	Score 1001;	DB 12;	Length 20489;					
	Best Local Similarity	100.0%;	Pred. No. 3.9e-231;							
	Matches 1001;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
QY	1	GCCCTGTGAGTGGAGAAAGGGCCCGAGAACTCAGAGATGGGACAGATACATCCTG	60							
DB	10500	GCCCTGTGGAGTGGAGAAAGGGGCCCGAAACTCAGAGATGGGACAGATACATCCTG	10559							
QY	61	AGGCAGAGGGGACACAGTGTGAGCTGCAGATCTGTGGCTGGCCATGGCGGACGCCGG	120							
DB	10560	AGGCAGAGGGGACCAAGTGTGAGCTGCAGATCTGTGGCTGGCCATGGCGGACGCCGG	10619							
QY	121	GAGTACTTGTGTGTGTCGGGCGCAGGAGGACCTCAGCCACGCTCACCATCAGGGCTGTG	180							
DB	10620	GAGTACTTGTGTGTGTCGGGCGCAGGAGGACCTCAGCCACGCTCACCATCAGGGCTGTG	10679							
QY	181	CTCTCCAGGTTCTAGAAGATGTGAAAAACCGAGAGGCCAGAGAGGGGCCACGGCTGTG	240							
DB	10680	CTCTCCAGGTTCTAGAAGATGTGAAAAACCGAGAGGCCAGAGAGGGGCCACGGCTGTG	10739							
QY	241	CTGCAGTGTGAGCTGAACAGTGCAGCCCTGTGGAGTGCAGAAAGGGGTCGAGACCTC	300							
DB	10740	CTGCAGTGTGAGCTGAACAGTGCAGCCCTGTGGAGTGCAGAAAGGGGTCGAGACCTC	10799							
QY	301	AGAGATGGGACAGATACACGCTCAGGCAGGACGGGACTAAATGTGAGCTGCAGATTGCT	360							
DB	10800	AGAGATGGGACAGATACACGCTCAGGCAGGACGGGACTAAATGTGAGCTGCAGATTGCT	10859							
QY	361	GGCCTGGCCATGGCAGACA CTGGGGAGTACTCTGTGCTGTGTGGGAGAGGACCTCG	420							
DB	10860	GGCCTGGCCATGGCAGACA CTGGGGAGTACTCTGTGCTGTGTGGGAGAGGACCTCG	10919							
QY	421	GCTATGCTCACGCTCAGGGCTCTACCCATCAAGTTCACAGAGGGTCTGAGGAACGAAG	480							
DB	10920	GCTATGCTCACGCTCAGGGCTCTACCCATCAAGTTCACAGAGGGTCTGAGGAACGAAG	10979							
QY	481	GCCACGAAGGGGCAACAGCCGTGCTGCGGTGTGAGCTGAGCAAGATGGCCCCGTGAG	540							
DB	10980	GCCACGAAGGGGCAACAGCCGTGCTGCGGTGTGAGCTGAGCAAGATGGCCCCGTGAG	11039							
QY	541	TGTTGGNAGGGGCATGAGACCTTCAGAGATGGAGACAGACAGGCTGAGGACGAGG	600							
DB	11040	TGTTGGNAGGGGCATGAGACCTTCAGAGATGGAGACAGACAGGCTGAGGACGAGG	11099							
QY	601	GCCAGGTGTGAGCTGCAGATCCGCGCCCTCTGTGGCAGAGGACGCTGGGGAGTACCTGTC	660							
DB	11100	GCCAGGTGTGAGCTGCAGATCCGCGCCCTCTGTGGCAGAGGACGCTGGGGAGTACCTGTC	11159							
QY	661	ATGTGCGGGAAGGAGGACCTCAGCCATGCTCACCGTCAGGGCCATGCCTTCCAAGTTC	720							
DB	11160	ATGTGCGGGAAGGAGGACCTCAGCCATGCTCACCGTCAGGGCCATGCCTTCCAAGTTC	11219							
QY	721	ATAGAGGCTCTGAGGAATGAAAGGCCACAGAAAGGGGATGAGACCTTCAGAGATGGGGAC	780							
DB	11220	ATAGAGGCTCTGAGGAATGAAAGGCCACAGAAAGGGGATGAGACCTTCAGAGATGGGGAC	11279							
QY	781	CTGAGCAAGCGGCCACCGGTGGAGTGGAGAAAGGGGCATGAGACCTTCAGAGATGGGGAC	840							
DB	11280	CTGAGCAAGCGGCCACCGGTGGAGTGGAGAAAGGGGCATGAGACCTTCAGAGATGGGGAC	11339							
QY	841	AGACACAGCCTGAGGCAGGAGGAGGCTCCAGGTGTGAGCTGCAGATCCGTGGCCTGGCTGTG	900							
DB	11340	AGACACAGCCTGAGGCAGGAGGAGGCTCCAGGTGTGAGCTGCAGATCCGTGGCCTGGCTGTG	11399							

agents for preventing a disease or condition associated with an aberrant or unwanted 59079 or 12599 activity in a subject, including cardiovascular diseases such as heart failure, and myocardial infarction; disorders involving blood vessels such as atherosclerosis, and Kaposi's sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia, Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders such as cancer; and protein kinase disorders such as autoimmune disorders, diabetes mellitus, psoriasis, inflammatory bowel disease, rheumatoid arthritis, and multiple sclerosis (many examples of diseases and disorders are included in the specification). The kinases, their encoding nucleic acids and antibodies are useful in screening assays, detection assays (e.g. forensic biology), and predictive medicine (e.g. diagnostic assays, prognostic assays, and monitoring clinical trials and pharmacogenomics). The kinases and their encoding nucleic acids are useful as query sequences to perform a search against public databases to identify other family members or related sequences. The present sequence encodes the kinase 12599

SQ Sequence 24120 BP; 4719 A; 7265 C; 8278 G; 3858 T; 0 U; 0 Other;

Query Match 100.0%; Score 1001; DB 8; Length 24120;
 Best Local Similarity 100.0%; Pred. No. 4.1e-231;
 Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCCCTGTGGAGTGGAGGAGCGGCCCGAGAACCTTCAGAGATGGGACAGATACATCCTG 60
 DB 10500 GCCCTGTGGAGTGGAGGAGCGGCCCGAGAACCTTCAGAGATGGGACAGATACATCCTG 10559

QY 61 AGGCAGGAGGGGACCCAGGTGTGAGCTGCAGATCTGTGGCCTGGCCATCGCGGAGCGCCGG 120
 DB 10560 AGGCAGGAGGGGACCCAGGTGTGAGCTGCAGATCTGTGGCCTGGCCATCGCGGAGCGCCGG 10619

QY 121 GAGTACTTGTGTGTGGGGCAGGAGAGGACCTCAGCCAGCTCACCATCAGGGCTCTG 180
 DB 10620 GAGTACTTGTGTGTGGGGCAGGAGAGGACCTCAGCCAGCTCACCATCAGGGCTCTG 10679

QY 181 CCTGCCAGTTCTATGAGATGTGAAGAACAGGAGGCCAGAGAGGGGCCACCGCTGTG 240
 DB 10680 CCTGCCAGTTCTATGAGATGTGAAGAACAGGAGGCCAGAGAGGGGCCACCGCTGTG 10739

QY 241 CTGCAGTGTGAGCTGAACAGTGCAGCCCTCTGGAGTGGAGAAAGGGGTCTGAGACCCCTC 300
 DB 10740 CTGCAGTGTGAGCTGAACAGTGCAGCCCTCTGGAGTGGAGAAAGGGGTCTGAGACCCCTC 10799

QY 301 AGAGATGGGACAGATACAGCTCAGGAGGACCGGACTAAATGTGAGCTGCAGATTCTG 360
 DB 10800 AGAGATGGGACAGATACAGCTCAGGAGGACCGGACTAAATGTGAGCTGCAGATTCTG 10859

QY 361 GGCCTGGCCATGGCGAGACACTGGGGAGTACTCTGCTGCTGGGGCAGGAGGACCTCG 420
 DB 10860 GGCCTGGCCATGGCGAGACACTGGGGAGTACTCTGCTGCTGGGGCAGGAGGACCTCG 10919

QY 421 GCTATGCTCAACGCTCAGGGCTCTACCCATCAAGTTTCACAGAGGGTCTCAGGAACGAGAG 480
 DB 10920 GCTATGCTCAACGCTCAGGGCTCTACCCATCAAGTTTCACAGAGGGTCTCAGGAACGAGAG 10979

QY 481 GCCACAGAGGGGCAACACCGCTGTGCGGTGTGAGCTGAGCAAGATGCCCCCGTGGAG 540
 DB 10980 GCCACAGAGGGGCAACACCGCTGTGCGGTGTGAGCTGAGCAAGATGCCCCCGTGGAG 11039

QY 541 TGGTGGAGGGGCATGAGACCTTCAGATGTGAGACAGACAGCCTCAGGACGAGACGGG 600
 DB 11040 TGGTGGAGGGGCATGAGACCTTCAGATGTGAGACAGACAGCCTCAGGACGAGACGGG 11099

QY 601 GCCAGGTGTGAGTGCAGATCGCGGCTCTGTGGCAGAGGAGCGCTGGGGAGTACCTGTGC 660
 DB 11100 GCCAGGTGTGAGTGCAGATCGCGGCTCTGTGGCAGAGGAGCGCTGGGGAGTACCTGTGC 11159

QY 661 ATGTGCGGGAAGGAGGAGGACCTTCAGCCATGTCTCAGCGTCAGGGCCATGCTTCCAAAGTTC 720
 DB 11160 ATGTGCGGGAAGGAGGAGGACCTTCAGCCATGTCTCAGCGTCAGGGCCATGCTTCCAAAGTTC 11219

QY 721 ATAGAGGCTCTGAGGAATGAAGAGGCCACAGAGGGGGACAGGGCCACGCTGTGTGTGAG 780

DB 11220 ATAGAGGGTCTGAGGANTGAGAGGCCACAGAGGGGACACGGCCACGCTGTGTGTGAG 11279
 QY 781 CTGAGCAAGCGCGCACCGGTGGAGTGTGAGGAGGAGGGGATGAGCCCTCAGAGATGGGGAC 840
 DB 11280 CTGAGCAAGCGCGCACCGGTGGAGTGTGAGGAGGAGGGGATGAGCCCTCAGAGATGGGGAC 11339
 QY 841 AGACACAGCTTGAGGACGAGGACGGGTCCAGGTGTGAGCTGCAGATCCGCTGGCCTGGCTGTG 900
 DB 11340 AGACACAGCTTGAGGACGAGGACGGGTCCAGGTGTGAGCTGCAGATCCGCTGGCCTGGCTGTG 11399
 QY 901 GTGATGTCGCGGGAGTACTCTGTGTGTGTGCGGGCAGGAGGAGCCTCAGCCACACTCACT 960
 DB 11400 GTGATGTCGCGGGAGTACTCTGTGTGTGTGCGGGCAGGAGGAGCCTCAGCCACACTCACT 11459
 QY 961 GTGAGGCGCTGCTGCTGCAGATTCATAGAGATGTGAAAAA 1001
 DB 11460 GTGAGGCGCTGCTGCTGCAGATTCATAGAGATGTGAAAAA 11500

RESULT 3
 ABZ24581

ID ABZ24581 standard; cDNA; 7564 BP.

XX AC ABZ24581;
 XX 31-MAR-2003 (first entry)
 DE Human cell adhesion and extracellular matrix protein 4 cDNA.
 KW Cell adhesion and extracellular matrix protein 4; CADECM-4; human;
 KW anti-HIV; virucide; antiallergic; antiinflammatory; antianaemic;
 KW antiparkinsonian; nootropic; anticonvulsant; antinfertility;
 KW antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;
 KW cytosatic; hepatotropic; dermatological; antidiabetic; nephrotropic;
 KW angiot; thymimetic; neuroprotective; osteopathic; antiarthritic;
 KW antiparasitic; antelmintic; antiprosoratic; utropathic; ophthalmological;
 KW antineumatic; haemostatic; antibacterial; protozoacide; fungicide;
 KW gynaecological; titin; gene therapy; gene; ss.

Homo sapiens.

Key Location/Qualifiers
 CDS 241..7227
 /*tag= a
 /product= "Human CADECM-4"

WO200288322-A2.

07-NOV-2002.

01-MAY-2002; 2002WO-US013874.

02-MAY-2001; 2001US-0288290P.

21-MAY-2001; 2001US-0292468P.

15-JUN-2001; 2001US-0298616P.

28-JUN-2001; 2001US-0301672P.

04-JAN-2002; 2002US-0345008P.

(INCY-) INCYTE GENOMICS INC.

Yue H, Lee EA, Duggan BM, Thangavelu K, Honchell CD, Ding L;

Hillman J, Baughn MR, Kallik DA, Lee S, Warren BA, Xu Y, Tran UK;

Lal PG, Thornton M, Hafalia AJA, Yao MG, Nguyen DB, Gandhi AR;

Khan FA, Walhia NK, Griffin JA, Chinn AM, Elliott VS, Ramkumar J;

Arvizu CS, Forsythe JJ;

WPI; 2003-167112/16.

P-PSDB; ABP58227.

New human cell adhesion and extracellular matrix proteins, useful for diagnosing, treating or preventing autoimmune or inflammatory disorder (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,

PT	cancer or hepatitis.	
XX	Claim 12; Page 168-170; 178pp; English.	
PS	The present sequence is that of Incyte clone 7326129CB1 cDNA encoding	
CC	human cell adhesion and extracellular matrix protein 4 (CADECM-4). A	
CC	representative cDNA library for the full-length polynucleotide is	
CC	MUSLTPR02, constructed from the muscle tissue RNA of a Caucasian adult	
CC	man. Homology data suggest the encoded protein to a titin muscle protein.	
CC	The invention provides CADECM-1 to -11 polypeptides (see ABP58224-34) and	
CC	polynucleotides (see ABZ24578-88), expression vectors, host cells,	
CC	antibodies, agonists and antagonists. These are useful for diagnosing,	
CC	treatment or preventing disorders associated with aberrant expression of	
CC	CADECM, particularly cell proliferative disorders (e.g. arteriosclerosis,	
CC	atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal	
CC	haemoglobinuria, polycythaemia vera, psoriasis, primary	
CC	thrombocytopaenia or cancer), developmental disorders (e.g. renal	
CC	tubular acidosis, anaemia or mental retardation), neurological disorders	
CC	(e.g. Alzheimer's disease, Parkinson's disease or epilepsy), reproductive	
CC	disorders (e.g. infertility or a disruption in the menstrual cycle), or	
CC	autoimmune/inflammatory disorders (e.g. AIDS, allergy, asthma, autoimmune	
CC	thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,	
CC	glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease,	
CC	Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis,	
CC	osteoarthritis, osteoporosis, pancreatitis, Reiter's syndrome, rheumatoid	
CC	arthritis, Sjogren's syndrome, uveitis), or viral, bacterial, fungal,	
CC	parasitic, protozoal or helminthic infections	
XX		
SQ	Sequence 7564 BP; 1594 A; 1999 C; 2711 G; 1260 T; 0 U; 0 Other;	
	Query Match 99.7%; Score 997.8; DB 8; Length 7564;	
	Best Local Similarity 99.8%; Pred. No. 1.8e-230;	
	Matches 999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;	
Qy	1 GCCCTGTGGAGTGGAGAGGGGCCCGAGAACCTCAGAGATGGGACAGATACATCTG 60	
Db	4546 GCCCTGTGGAGTGGAGAGGGGCCCGAGAACCTCAGAGATGGGACAGATACATCTG 4605	
Qy	61 AGGCAGAGGGGACAGGTGTGAGCTGAGCTGAGATCTGTGGCTTGGCCATGGCGGCGGG 120	
Db	4606 AGGCAGAGGGGACAGGTGTGAGCTGAGATCTGTGGCTTGGCCATGGCGGCGGG 4665	
Qy	121 GAGTACTTGTGTGTGCGGACAGAGAGGACCTCAGCAGCTCACCATCAGGGCTCTG 180	
Db	4666 GAGTACTTGTGTGTGCGGACAGAGAGGACCTCAGCAGCTCACCATCAGGGCTCTG 4725	
Qy	181 CCTGCCAGTTCTAGAGATGTGAAACACAGAGGCGCAGAGAGGGGCAACGGCTGTG 240	
Db	4726 CCTGCCAGTTCTAGAGATGTGAAACACAGAGGCGCAGAGAGGGGCAACGGCTGTG 4785	
Qy	241 CTGCAGTGTGAGCTGAACAGTGACGCCCTGTGAGTGGAGAAAGGGGTCTGACACCTC 300	
Db	4786 CTGCAGTGTGAGCTGAACAGTGACGCCCTGTGAGTGGAGAAAGGGGTCTGACACCTC 4845	
Qy	301 AGAGATGGGACAGATACAGCTCAGGACGAGGACCTAAATGTGAGCTGCAGATTCTG 360	
Db	4846 AGAGATGGGACAGATACAGCTCAGGACGAGGACCTAAATGTGAGCTGCAGATTCTG 4905	
Qy	361 GGCTTGGCCATGGCAGACACTGGGGAGTACTCTGCTGCTGGCGGACGAGAGACCTCG 420	
Db	4906 GGCTTGGCCATGGCAGACACTGGGGAGTACTCTGCTGCTGGCGGACGAGAGACCTCG 4965	
Qy	421 GCTATGCTCACCGTCAAGGCTCTACCCATCAAGTTTCAGAGGGTCTGAGGAACGAGAG 480	
Db	4966 GCTATGCTCACCGTCAAGGCTCTACCCATCAAGTTTCAGAGGGTCTGAGGAACGAGAG 5025	
Qy	481 GCCACAGAAGGGGCAACAGCCGCTGCTGGGTGTGAGCTGAGCAAGATGGCCCCCGTGGAG 540	
Db	5026 GCCACAGAAGGGGCAACAGCCGCTGCTGGGTGTGAGCTGAGCAAGATGGCCCCCGTGGAG 5085	
Qy	541 TGGTGGAGGGGCTAGACACCTCAGAGATGGAGACAGACACAGCCTGAGGACGAGCGG 600	
Db	5086 TGGTGGAGGGGCTAGACACCTCAGAGATGGAGACAGACACAGCCTGAGGACGAGCGG 5145	
Qy	601 GCCAGGTGTGAGCTGCAGATCCGGGCTCGTGGCAGAGACGCTGGGGAGTACTCTGTC 660	
Db	5146 GCCAGGTGTGAGCTGCAGATCCGGGCTCGTGGCAGAGACGCTGGGGAGTACTCTGTC 5205	
Qy	661 ATGTGCGGGGAAGAGAGAGGACCTCAGCCATCTCACCTCAGGGCCATGCCCTTCCAAGTTC 720	
Db	5206 ATGTGCGGGGAAGAGAGAGGACCTCAGCCATCTCACCTCAGGGCCATGCCCTTCCAAGTTC 5265	
Qy	721 ATAGAGGCTCTGAGGAATGAAGAGGCCACAGAGGGGACACGGCCACGCTGTGTTGTGAG 780	
Db	5266 ATAGAGGCTCTGAGGAATGAAGAGGGGCCACAGAGGGGACACGGCCACGCTGTGTTGTGAG 5325	
Qy	781 CTGAGCAAGGCGGCACCGTGGAGTGGAGGAAGGGGCATCAGACCTCAGAGATGGGGAC 840	
Db	5326 CTGAGCAAGGCGGCACCGTGGAGTGGAGGAAGGGGCATCAGACCTCAGAGATGGGGAC 5385	
Qy	841 AGACACAGCCTGAGGGCAGGACGGTCCAGTGTGAGCTGCAGATCCCTGCGCTGGCTGTG 900	
Db	5386 AGACACAGCCTGAGGGCAGGACGGTCCAGTGTGAGCTGCAGATCCCTGCGCTGGCTGTG 5445	
Qy	901 GTGATGCCGGGAGTACTCTGTGTGTGCGGGCAGGAGGACCTCAGCCACTCCTC 960	
Db	5446 GTGATGCCGGGAGTACTCTGTGTGTGCGGGCAGGAGGACCTCAGCCACTCCTC 5505	
Qy	961 GTACGGGCTCTGCTGCCAGATTTCATAGAAGATGTGAAAAA 1001	
Db	5506 GTACGGGCTCTGCTGCCAGATTTCATAGAAGATGTGAAAAA 5546	
RESULT 4		
ABV99363	ABV99363 standard; DNA; 14061 BP.	
ID	ABV99363 standard; DNA; 14061 BP.	
XX		
AC	ABV99363;	
XX		
DT	27-JAN-2003 (first entry)	
XX		
DE	Human NOV13b coding sequence.	
XX		
KW	Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;	
KW	antiinflammatory; cardiac; haemostatic; neuroprotective; anorectic;	
KW	neotrophic; immunosuppressive; osteopathic; antiparkinsonian; cancer;	
KW	infertility; cerebroprotective; gene therapy; NOVX; NOV; fertility;	
KW	metabolic disorder; diabetes; obesity; infectious disease; anorexia;	
KW	neurodegenerative disease; Alzheimer's disease; Parkinson's disease;	
KW	immune disorder; haematopoietic disorder; cardiovascular disorder;	
KW	bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;	
KW	metabolic syndrome X; wasting disorder; cell differentiation; gene;	
KW	cell proliferation; haematopoiesis; wound healing; angiogenesis; ds.	
XX		
OS	Homo sapiens.	
XX		
FN	WO200272771-A2.	
XX		
PD	19-SEP-2002.	
XX		
PF	08-MAR-2002; 2002WO-US007288.	
XX		
PR	08-MAR-2001; 2001US-0274101P.	
PR	08-MAR-2001; 2001US-0274194P.	
PR	08-MAR-2001; 2001US-0274281P.	
PR	08-MAR-2001; 2001US-0274322P.	
PR	09-MAR-2001; 2001US-0274849P.	
PR	12-MAR-2001; 2001US-0275235P.	
PR	13-MAR-2001; 2001US-0275578P.	
PR	13-MAR-2001; 2001US-0275579P.	
PR	13-MAR-2001; 2001US-0275601P.	
PR	14-MAR-2001; 2001US-0276000P.	
PR	16-MAR-2001; 2001US-0276776P.	
PR	19-MAR-2001; 2001US-0276994P.	
PR	20-MAR-2001; 2001US-0277239P.	
PR	20-MAR-2001; 2001US-0277321P.	

PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277791P.
PR 22-MAR-2001; 2001US-0277833P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278894P.
PR 27-MAR-2001; 2001US-0278999P.
PR 27-MAR-2001; 2001US-0279038P.
PR 28-MAR-2001; 2001US-0279344P.
PR 30-MAR-2001; 2001US-0279995P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280909P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 30-APR-2001; 2001US-0287424P.
PR 03-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 03-MAY-2001; 2001US-0288528P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291109P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294689P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318462P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 27-SEP-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 14-NOV-2001; 2001US-0333272P.
PR 21-NOV-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337428P.
PR 03-DEC-2001; 2001US-0338092P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
PI Boldog FL, Li L, Zerhusen BD, Tchernev VT, Gangolli EA, Vernet CM;
PI Pena CEA, Burgess CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
PI Voss EE, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
PI Taupier RJ, Padigar M, Shency SG, Kekuda R, Gusev VY, Pochart PF;
PI Zhong M;
XX
XX WPI, 2002-732824/79.
DR P-PSDB; ABP70085.
DR
XX
XX New NOVX polypeptides and polynucleotides, useful for preventing,
PT diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
PT Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
PT disorders, and asthma.
XX
ES Claim 16; Page 138-142; 619pp; English.
XX
XX The present invention relates to new isolated proteins (NOVX) and their
CC coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
CC any number from 1 to 48. The NOVX proteins and coding sequences are
CC useful in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, preferably a NOVX-associated disorder.
CC The NOVX coding sequences and proteins are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious diseases, anorexia, cancer-associated cachexia,
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
CC disease, immune disorders, hematopoietic disorders, cardiovascular
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, hematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods
XX
SQ Sequence 14061 BP; 2854 A; 4029 C; 4839 G; 2339 T; 0 U; 0 Other;
Query Match 99.7%; Score 997.8; DB 6; Length 14061;
Best Local Similarity 99.8%; Pred. No. 2.1e-230;
Matches 999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCCCTCTGTGAGTGGAGAGGGGCCCGAGAACCTTCAGAGATGGGGACAGATACATCTCTG 60
DB 2661 GCCCTCTGTGAGTGGAGAGGGGCCCGAGAACCTTCAGAGATGGGGACAGATACATCTCTG 2720
QY 61 AGGCAGAGGGGACCAAGGTGTGAGCTGCAGATCTGTGTGGCTGGCCATCGCGAGACCCGGG 120
DB 2721 AGGCAGAGGGGACCAAGGTGTGAGCTGCAGATCTGTGTGGCTGGCCATCGCGAGACCCGGG 2780
QY 121 GAGTACTTGTGTGTGTGGGGCAGGAGGAGGACCTTCAGCCACGCTCACCATCAGGGCTCTG 180
DB 2781 GAGTACTTGTGTGTGTGGGGCAGGAGGAGGACCTTCAGCCACGCTCACCATCAGGGCTCTG 2840
QY 181 CCTGCCAGTTTCATAGAAGATGTGAAAAACAGAGGCGCAGAGAGGGGCCACCGCTGTG 240
DB 2841 CCTGCCAGTTTCATAGAAGATGTGAAAAACAGAGGCGCAGAGAGGGGCCACCGCTGTG 2900
QY 241 CTGCAGTGTGAGCTGAACAGTGCAGCCCTCTGTGAGTGGAGAGAAAGGGGTCTGAGACCTTC 300
DB 2901 CTGCAGTGTGAGCTGAACAGTGCAGCCCTCTGTGAGTGGAGAGAAAGGGGTCTGAGACCTTC 2960
QY 301 AGAGATGGGGACAGATACAGCCTTCAGGCAGGACGGGACTTAATGTGAGCTGCAGATTCTG 360
DB 2961 AGAGATGGGGACAGATACAGCCTTCAGGCAGGACGGGACTTAATGTGAGCTGCAGATTCTG 3020
QY 361 GGCTTGGCCATGGCAGACACTTGGGGAGTACTCTGCGGTGTGGGGCAGGAGGACCTCG 420
DB 3021 GGCTTGGCCATGGCAGACACTTGGGGAGTACTCTGCGGTGTGGGGCAGGAGGACCTCG 3080
QY 421 GCTATGCTCACCGTCAGGGCTCTACCCATCAAGTTTCAGAGGGTCTCAGGAAACGAAGAG 480
DB 3081 GCTATGCTCACCGTCAGGGCTCTACCCATCAAGTTTCAGAGGGTCTCAGGAAACGAAGAG 3140
QY 481 GCCCAGAGGGGCAACAGCCGCTCTGCGGTGTGAGCTGACCAAGATGGCCCCCGTGGAG 540
DB 3141 GCCCAGAGGGGCAACAGCCGCTCTGCGGTGTGAGCTGACCAAGATGGCCCCCGTGGAG 3200
QY 541 TGGTGGAGGGGCATGAGACCTTCAGAGATGGAGACAGACAGCCTTCAGGCAGGACGGG 600
DB 3201 TGGTGGAGGGGCATGAGACCTTCAGAGATGGAGACAGACAGCCTTCAGGCAGGACGGG 3260
QY 601 GCCAGGTGTGAGCTGCAGATCCGGGCTCTGTGGCAGAGGACGCTGGGGAGTACTCTGTGC 660
DB 3261 GCCAGGTGTGAGCTGCAGATCCGGGCTCTGTGGCAGAGGACGCTGGGGAGTACTCTGTGC 3320
QY 661 ATGTGCGGGAGGAGGAGGACCTTCAGCCATGCTCACCCTCAGGGCCATCCCTTCCAAGTTC 720
DB 3321 ATGTGCGGGAGGAGGAGGACCTTCAGCCATGCTCACCCTCAGGGCCATCCCTTCCAAGTTC 3380
QY 721 ATAGAGGTCTGAGGAATGAAGAGGCCACAGAAAGGGGACAGCCGCTGTGGTGTGAG 780
DB 3381 ATAGAGGTCTGAGGAATGAAGAGGCCACAGAAAGGGGACAGCCGCTGTGGTGTGAG 3440

CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods

XX SQ Sequence 14109 BP; 2862 A; 4045 C; 4854 G; 2348 T; 0 U; 0 Other;

Query Match 99.7%; Score 997.8; DB 6; Length 14109;
Best Local Similarity 99.8%; Pred. No. 2.1e-230;
Matches 999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCTGTGGAGTGGAGAGGGGCCGAGAACCTCAGAGATGGGGACAGATACATCTCTG 60
Db 2661 GCCCTGTGGAGTGGAGAGGGGCCGAGAACCTCAGAGATGGGGACAGATACATCTCTG 2720

Qy 61 AGGAGAGGGGACACAGTGTGAGCTGCAGATCTGTGGCTGGCCATGGCGGACGCCGGG 120

Db 2721 AGGAGAGGGGACACAGTGTGAGCTGCAGATCTGTGGCTGGCCATGGCGGACGCCGGG 2780

Qy 121 GAGTACTTGTGTGTGGCGGACGAGAGGACCTCAGCCACGCTCACCATCAGGGCTCTG 180

Db 2781 GAGTACTTGTGTGTGGCGGACGAGAGGACCTCAGCCACGCTCACCATCAGGGCTCTG 2840

Qy 181 CCTGCCAGGTTTCATAGAGATGTGAAAAACAGAGAGGCCAGAGAGGGGCCACGGCTGTG 240

Db 2841 CCTGCCAGGTTTCATAGAGATGTGAAAAACAGAGAGGCCAGAGAGGGGCCACGGCTGTG 2900

Qy 241 CTCAGTGTGAGCTGAACAGTGCAGCCCTCTGTGGAGTGGAGAAAGGGGTCTGAGACCTC 300

Db 2901 CTCAGTGTGAGCTGAACAGTGCAGCCCTCTGTGGAGTGGAGAAAGGGGTCTGAGACCTC 2960

Qy 301 AGAGATGGGACAGATACAGCTGAGGACGAGGCGGACTTAATGTGAGCTCAGATTCTG 360

Db 2961 AGAGATGGGACAGATACAGCTGAGGACGAGGCGGACTTAATGTGAGCTCAGATTCTG 3020

Qy 361 GGCCTGGCCATGGGACACACTGGGAGTACTCTGCTGTGGCGGACGAGAGGACCTCG 420

Db 3021 GGCCTGGCCATGGGACACACTGGGAGTACTCTGCTGTGGCGGACGAGAGGACCTCG 3080

Qy 421 GCTATGCTCACCGTCAAGGCTCTACCCATCAAGTTTCACAGAGGGTCTCAGGAACGAG 480

Db 3081 GCTATGCTCACCGTCAAGGCTCTACCCATCAAGTTTCACAGAGGGTCTCAGGAACGAG 3140

Qy 481 GCCACAGAGGGGCAACAGCCGTCTGGGTGTGAGCTGAGCAAGATGGCCCCCGTGGAG 540

Db 3141 GCCACAGAGGGGCAACAGCCGTCTGGGTGTGAGCTGAGCAAGATGGCCCCCGTGGAG 3200

Qy 541 TGGTGAAGGGCATGAGACCTCAGAGATGGAGACAGACAGCCTCAGGCAGGACGGG 600

Db 3201 TGGTGAAGGGCATGAGACCTCAGAGATGGAGACAGACAGCCTCAGGCAGGACGGG 3260

Qy 601 GCCAGGTGTGAGCTCAGATCCGGGCTCTGGTGGCAGAGGACGCTGGGGAGTACCTGTGC 660

Db 3261 GCCAGGTGTGAGCTCAGATCCGGGCTCTGGTGGCAGAGGACGCTGGGGAGTACCTGTGC 3320

Qy 661 ATGTCCGGGAGGAGAGACCTCAGCATGCTCACCCTCAGGGCCATGCTTCCAAGTTC 720

Db 3321 ATGTCCGGGAGGAGAGACCTCAGCATGCTCACCCTCAGGGCCATGCTTCCAAGTTC 3380

Qy 721 ATAGAGGGTCTGAGCAATGAGAGGCCACAGAGGGGACAGCGCCACGCTGTGTGTGAG 780

Db 3381 ATAGAGGGTCTGAGCAATGAGAGGGGACAGAGGGGACAGCGCCACGCTGTGTGTGAG 3440

Qy 781 CTGAGCAAGGGGCGACCCGGTGGAGTGGAGAGGGGCGATGAGACCTCAGAGATGGGAC 840

Db 3441 CTGAGCAAGGGGCGACCCGGTGGAGTGGAGAGGGGCGATGAGACCTCAGAGATGGGAC 3500

Qy 841 AGACACAGCTTGAAGCAGGACCGGTTCAGGTGTGAGCTGCAGATCCGTGGCTGGCTGTG 900

Db 3501 AGACACAGCTTGAAGCAGGACCGGTTCAGGTGTGAGCTGCAGATCCGTGGCTGGCTGTG 3560

Qy 901 GTGGATCCGGGGAGTACTCTGTGTGTGTCGGGACGAGAGACCTCAGCCACACTCACT 960

Db 3561 GTGGATCCGGGGAGTACTCTGTGTGTGTCGGGACGAGAGACCTCAGCCACACTCACT 3620

Qy 961 GTGAGGCCCCCTGCTGCCAGATTTCATAGAGATGTGAAAA 1001
Db 3621 GTGAGGCCCCCTGCTGCCAGATTTCATAGAGATGTGAAAA 3661

RESULT 6

AAF64470/C

ID AAF64470 standard; cDNA; 352 BP.

XX AC AAF64470;

XX DT 09-APR-2001 (first entry)

XX Novel human polynucleotide, SEQ ID NO: 226.

XX Human; cytostatic; gene therapy; colon cancer; prostate cancer;
XX breast cancer; lung cancer; cancer detection; ss.

XX OS Homo sapiens.

XX FN WO200102568-A2.

XX PD 11-JAN-2001.

XX PF 30-JUN-2000; 2000WO-US018374.

XX PR 02-JUL-1999; 99US-0142310P.

XX PR 02-JUL-1999; 99US-0142311P.

XX (CHIR) CHIRON CORP.

XX (HYSE-) HYSEQ INC.

XX Williams LT, Escobedo J, Innis MA, Garcia PD, Klinger J;
XX Kassar A, Reinhard C, Randazzo F, Kennedy GC, Pot D, Lamson G;
XX Drmanac R, Crkenjakov R, Drmanac S, Dickson M, Labat I;
XX Leshkowitz D, Kita D, Garcia V, Jones LW, Strache-Crain B;

XX WPI; 2001-091805/10.

XX Library of polynucleotides for diagnosing a cancerous state of a
XX mammalian cell and detecting cancer, particularly of the colon or
XX prostate, comprises 3351 human polynucleotide sequences.

XX Claim 9; Page 576; 1046pp; English.

XX The present sequence is one of 3351 sequences in a library of human
XX polynucleotides. The library is used to detect differentially expressed
XX genes correlated with a cancerous state of a mammalian cell and can
XX detect colon, prostate, breast and lung cancer. The library can be used
XX to produce probes for detection of mRNA and to produce additional copies
XX of the polynucleotides. The probes can be used for chromosome mapping of
XX the polynucleotide and for detection of transcription levels. Ribozymes
XX or antisense oligonucleotides can be generated. The polynucleotides and
XX their gene products are used as genetic or biochemical markers (e.g. in
XX blood or tissues) that will detect the earliest changes along the
XX carcinogenesis pathway and/or monitor the efficacy of therapies and
XX preventive interventions. The polynucleotides, polypeptides and
XX antibodies against them can be used in pharmaceutical compositions to
XX treat the cancers and proliferative disorders such as neoplasia,
XX dysplasia and hyperplasia

SQ Sequence 352 BP; 56 A; 138 C; 84 G; 73 T; 0 U; 1 Other;

Query Match 34.3%; Score 343.4; DB 5; Length 352;

Best Local Similarity 99.7%; Pred. No. 5.1e-73;

Matches 344; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 631 GTGGCAGAGGACGCTGGGGAGTACCTGTGCAATGTGCGGGAAGGAGGACCTCAGCCATG 690

Db 351 GTGGCAGAGGACGCTGGGGAGTACCTGTGCAATGTGCGGGAAGGAGGACCTCAGCCATG 292

Qy 691 CTCACCGTCAGGGGCCATGCCCTTCCAAAGTTCTAGAGGCTCTGAGGAATGAGAGGCCACA 750

Db	291	CTCACCGTCAGGGCCATGCTTCCAAAGTTCATAGAGGCTCTGAGGAATGAAGAGGCCACA	232
Qy	751	GAAGGGGACACGGGCCACGCTGTGTGTGAGCTGAGCAAGGCGGCA	810
Db	231	GAAGGGGACACGGGCCACGCTGTGTGTGAGCTGAGCAAGGCGGCA	172
Qy	811	AAGGGGCATGAGACCCCTCAGAGATGGGACAGACACAGCCCTTAGGGCAGACCGGTCACAG	870
Db	171	AAGGGGCATGAGACCCCTCAGAGATGGGACAGACACAGCCCTTAGGGCAGAGTGGTCCAGG	112
Qy	871	TGTGAGCTCAGATCGGTGGCCTGCGTGTGTGTGATGCGCGGGAGTAGTACTCGTGTGTGTGC	930
Db	111	TGTGAGCTCAGATCCGTGGCCCTGCGCTGTGTGTGATGCCGGGAGTAGTACTCGTGTGTGTGC	52
Qy	931	GGCAGGAGAGGACCTTCAGCCACACTCACTGTGAGGGCCCTGCCT	975
Db	51	GGCAGGAGAGGACCTTCAGCCACACTCACTGTGAGGGCCCTGCCT	7

RESULT 7

ABQ61169
ID ABQ61169 standard; cDNA; 642 BP.

AC ABO61169:

DT 26-FEB-2003 (first entry)

DE Obscurin OBSCN gene encoding sequence.

Neuroprotective; immunomodulator; cancer; cytostatic; anti-inflammatory;
 KW gene therapy; nutritional supplement; wound; burn; ulcer;
 Alzheimer's disease; Huntington's disease; amyotrophic lateral sclerosis;
 KW autoimmune disorder; inflammation; vulvovaginitis; gene; ss.

Homo sapiens.

PN WO200231111-A2.

PD 18-APR-2002.

PF 11-OCT-2001; 2001WO-US027760.

PR 12-OCT-2000; 2000US-00687527.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Wehrman T, Drmanac RT;

DR WPI; 2002-426278/45.

DR N-PSDB; ABP43925.

New polypeptides and their encoded proteins, useful as nutritional sources or supplements, or in gene therapy, particularly for treating wounds, Alzheimer's disease, amyotrophic lateral sclerosis, cancer or inflammation.

PS Claim 1; SEQ ID # 382; 357pp + Sequence Listing; English.

The invention relates to 446 newly isolated polynucleotide sequences. The activity of polynucleotides of the invention may be described as, vulnary, neuroprotective, immunomodulator, cytostatic and anti-inflammatory. Compositions comprising nucleic acids of the invention are useful for treating a mammalian subject, or as nutritional sources or supplements. These are useful in gene therapy, particularly for treating wounds, burns or ulcers, Alzheimer's disease, Huntington's disease, amyotrophic lateral sclerosis, autoimmune disorders, cancer or inflammation. The nucleic acids and polypeptides are also useful in diagnostic and research methods. The sequences given in records ABQ6798-ABQ61233 represent polynucleotides of the invention. NOTE: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at

CC useful for treating a mammalian subject, or as nutritional sources or
 CC supplements. These are useful in gene therapy, particularly for treating
 CC wounds, burns or ulcers, Alzheimer's disease, Huntington's disease,
 CC amyotrophic lateral sclerosis, autoimmune disorders, cancer or
 CC inflammation. The nucleic acids and polypeptides are also useful in
 CC diagnostic and research methods. The sequences given in records AB060788-
 CC AB061233 represent polynucleotides of the invention. NOTE: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX
 SQ Sequence 707 BP; 154 A; 198 C; 213 G; 142 T; 0 U; 0 Other;
 Query Match 26.5%; Score 265; DB 6; Length 707;
 Best Local Similarity 100.0%; Pred. No. 4.7e-54;
 Matches 265; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 439 GCTCTACCCATCAAGTTTCACAGAGGGTCTGAGGAACGAGAGGCCACAGAGGGGCCAACA 498
 Db 227 GCTCTACCCATCAAGTTTCACAGAGGGTCTGAGGAACGAGAGGCCACAGAGGGGCCAACA 286

Qy 499 GCCGTGCTGCGGTGTGAGCTGACGACGACGATGCCCGCCCGTGGAGTGGTGGAGGGGCATGAG 558
 Db 287 GCCGTGCTGCGGTGTGAGCTGACGACGACGATGCCCGCCCGTGGAGTGGTGGAGGGGCATGAG 346

Qy 559 ACCCTCAGATGAGACAGACACAGCCTGAGGACGAGCGGGCCAGGTGTGAGCTGCAG 618
 Db 347 ACCCTCAGATGAGACAGACACAGCCTGAGGACGAGCGGGCCAGGTGTGAGCTGCAG 406

Qy 619 ATCCGCGCGCTCGTGGGACGAGCGCTGGGAGTACCTGTCATGTCGGGGAAGGAGG 678
 Db 407 ATCCGCGCGCTCGTGGGACGAGCGCTGGGAGTACCTGTCATGTCGGGGAAGGAGG 466

Qy 679 ACCTCAGCAGTGTCCACCGTCAGG 703
 Db 467 ACCTCAGCAGTGTCCACCGTCAGG 491

RESULT 9
 ABK99965
 ID ABK99965 standard; DNA; 2155 BP.
 AC ABK99965;
 XX
 XX 21-OCT-2002 (first entry)
 XX
 XX DNA encoding human secreted protein SCEP-48.
 XX
 KW Secreted protein; SCEP; human; cell proliferative disorder; cancer;
 KW keratosis; arteriosclerosis; atherosclerosis; cirrhosis; hepatitis;
 KW psoriasis; autoimmune disorder; inflammatory disorder; AIDS; arthritis;
 KW acquired immunodeficiency syndrome; adult respiratory distress syndrome;
 KW Addison's disease; allergy; asthma; osteoporosis; autoimmune thyroiditis;
 KW Crohn's disease; dermatitis; diabetes; Graves' disease; haemodialysis;
 KW glomerulonephritis; scleroderma; systemic lupus erythematosus; uveitis;
 KW systemic sclerosis; ulcerative colitis; infection; trauma; Pick disease;
 KW cardiovascular disorder; congestive heart failure; angina; epilepsy;
 KW hypertensive heart disease; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; amyotrophic lateral sclerosis; stroke; dementia;
 KW metabolic disorder; endocrine disorder; toxic myopathy; mental disorder;
 KW schizophrenic disorder; developmental disorder; anaemia; epilepsy;
 KW hypothyroidism; glaucoma; sensorineural hearing loss; cataract;
 KW transgenic animal; gene; da.
 XX
 OS Homo sapiens.
 XX
 PN W0200248337-A2.
 XX
 XX 20-JUN-2002.
 PD
 XX 12-DEC-2001; 2001WO-US048517.
 PF
 XX

PR 13-DEC-2000; 2000US-0255633P.
 PR 21-DEC-2000; 2000US-0257852P.
 PR 05-JAN-2001; 2001US-0260105P.
 PR 18-JAN-2001; 2001US-0262932P.
 PR 18-JAN-2001; 2001US-0263096P.
 PR 19-JAN-2001; 2001US-0263090P.
 PR 02-FEB-2001; 2001US-0265926P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 PI Griffin JA, Yao MG, Duggan BM, Yue H, Ding L, Lal PG, Lee EA;
 PI Ramkumar J, Thangavelu K, Xu Y, Lee S, Tang Y, Nguyen DB;
 PI Warren BA, Honcheill CD, Gietzen KJ, Baughn MR, Gandhi AR, Arvizu C;
 PI Walia NK, Lu Y, Elliott VW, Lu DAM, Hafalia AUA, Azimzai Y;
 PI Khan FA, Tran UK;
 XX
 DR WPI; 2002-583509/62.
 DR P-PSDB; ABG69668.
 XX
 PT Novel human secreted proteins and polynucleotides for diagnosing,
 PT preventing or treating cell proliferative, autoimmune/inflammatory,
 PT cardiovascular, neurological and developmental disorders.
 PS
 PS Claim 5; Page 230; 234pp; English.
 XX
 CC The invention describes an isolated polypeptide chosen from secreted
 CC proteins (I), SCEP 1-54. (I) and the polynucleotide encoding it (II) are
 CC useful for screening a compound for effectiveness as an agonist or
 CC antagonist of (I) or compound that alters expression of (II). (I), the
 CC identified agonist and antagonist are useful for treating a disease or
 CC condition associated altered expression of functional SCEP in a patient.
 CC An antibody specific to (I) is useful for detecting the presence of (I).
 CC Purifying (I) from a sample and for diagnosing a condition or disease
 CC associated with expression of SCEP in a subject or in a biological
 CC sample. (I) and (II) and modulators of (I) are useful for diagnosis,
 CC treatment and prevention of cell proliferative disorders (e.g. cancer,
 CC keratosis, arteriosclerosis, atherosclerosis, cirrhosis, hepatitis and
 CC psoriasis), autoimmune/inflammatory disorders (e.g. acquired
 CC immunodeficiency syndrome (AIDS), adult respiratory distress syndrome,
 CC Addison's disease, allergies, asthma, osteoporosis, autoimmune
 CC thyroiditis, Crohn's disease, dermatitis, diabetes, Graves' disease,
 CC glomerulonephritis, arthritis, scleroderma, systemic lupus erythematosus,
 CC systemic sclerosis, ulcerative colitis, haemodialysis, uveitis; viral,
 CC bacterial, fungal, parasitic, protozoal, helminthic infections and
 CC trauma), cardiovascular disorders (e.g. congestive heart failure, angina,
 CC hypertensive heart disease), neurological disorders (e.g. Alzheimer's and
 CC Pick disease, Parkinson disease, amyotrophic lateral sclerosis, epilepsy,
 CC stroke, Huntington's disease, multiple sclerosis, dementia, neuromuscular
 CC disorders, metabolic, endocrine and toxic myopathies, mental disorders,
 CC schizophrenic disorders, and developmental disorders (e.g. anaemia,
 CC epilepsy, hypothyroidism, glaucoma, sensorineural hearing loss and
 CC cataract). (II) is useful for creating transgenic animals to model human
 CC disease and to detect and quantify gene expression in biopsied tissues in
 CC which expression of SCEP is correlated with disease. This sequence
 CC encodes a human secreted protein (SCEP)
 XX
 SQ Sequence 2155 BP; 488 A; 530 C; 545 G; 592 T; 0 U; 0 Other;
 Query Match 23.9%; Score 238.8; DB 6; Length 2155;
 Best Local Similarity 59.8%; Pred. No. 1.3e-47;
 Matches 531; Conservative 0; Mismatches 282; Indels 95; Gaps 4;

Qy 175 GCTCTGCTGCCAGGTTTCATAGAGATGTGANAACAGGAGGCCAGAGAGGGGCCACG 234
 Db 205 GCCCTGCTGCCAGGTTTCACAGAGGGTCTGAGGAATGAGAGGCCATGAGGGGCCACA 264

Qy 235 GCTGTGCTGCAGTGTGAGCTGAACAGTGCAGCCCTCTGGAGTGGAGAAAGGGTCTGAG 294
 Db 265 GCCACACTGCAATGTGAGCTGAGCAAGGCAGCCCTCTGTGGAGTGGAGAAAGGGCTTGAG 324

Qy 295 ACCCTCAGATGGGACAGATACAGCTGAGGACGAGCGGACCTTAAATGTGAGCTGCAG 354
 Db 325 GCTCTCAGATGGGGACAATAATACAGCTGAGACAGCGGGCTGTGTGTGAGCTGCAG 384

QY	355	ATTTCGTGGCCTGGCCCATGGCAGACACTGGGAGTACTCGTGGCTGTGCGGGCAGGAGAGG	414
Db	385		
QY	415	ACCTCGGTATGCTCACCGTCAGGGCTCTACCCATCAAGTTTCACAGACGGGTC-----	466
Db	445		
QY	467	TGAGGAACGAAGAGCCACAGAGGGGGCAACGCCGTGCTGCGGTGTGAGCTGAGCAAGA	526
Db	505	TGGTGATACACCTCTCTGTGTACCAACCTTCGCCTCCAAATGTGGCACATCTCCTGTG	564
QY	527	TGGCCCCCGTGGAGTGTGGAAAGGGGCATGACCCCTCAGAG-----	568
Db	565	GAACCCCTGTGATTGTCTCTCTAACTTGGGGCCCTTAGGTATCTCTGCCCTCCCTT	624
QY	569	--ATGGAGACAGACACAGCCCTGAGGCAGAGCGGGCCAGGTGTGAGCTGCAGATCCCGGG	626
Db	625	CTCATCACTAGATGTCCTCTCTGCTCATGCAATGTCCTGTGTGTTCTCTGTGCCCTCTGGG	684
QY	627	CCTCGTGGCAGAGGACGCTGGGGA-----	650
Db	685	CCATTGGCTTCTCATTTGTTTATATATTGCTCATCTAGCTATGGTCCCTTGGTGTTGT	744
QY	651	-----CTACCTGTGCATGTGCGGAGAGGAGAGACCTCAGCCATGTCTCA-----	694
Db	745	GTGGCTCCTTATTGGTGTCCCATGTTCTGTCCGAAAAATCCTCCACAGACAGTCTGATGAT	804
QY	695	-----CCGTCAGGGCCATGCCTTCCAAGTTCATAGAGGCTCTGAGGAATG	739
Db	805	CAGTGACTGTTTGGTCTCTCTCAGCCCTGCCTGCCAGATTTCATAGAGGATATGAGAAACC	864
QY	740	AAGAGGCCACAGAGGGGACACGGCCACGCTGTGGTGTGAGCTGAGCAAGCGCCACCGG	799
Db	865	AGAAGGCCACAGAAAGGGGCTACAGTCAATTGTCAATGTAACTCAGAAAAGCGCGCCCCG	924
QY	800	TGGAGTGGAGGAAGGGGCATAGACCTCTCAGAGATGGGACAGACACAGCTGTAGGCAGG	859
Db	925	TGGAGTGGAGAAAGGGGCCAACACCTTCAAAGATGGGGAACAGGTACAGCCTGAAGCAGG	984
QY	860	ACGGGTCCAGGTGTGAGCTGCAGATCCGTGGCCCTGGCTGTGGTGGATGCCGGGAGGTACT	919
Db	985	ATGGGACCAAGTTGTGAGCTGCAGATTCTGTGSCCTGGTTCATAGCAGATGCTGGAGAACT	1044
QY	920	CGTGTGTGCGGGCAGAGAGGACCTCAGCCCACTCACTGTCAAGG	967
Db	1045	CGTGCATATGTGACAGAGAGGACCTCGGCCACGCTCACTGTCAAGG	1092

RESULT 10	
ABL58139	
ID	ABL58139 standard; cDNA; 898 BP.
XX	
XX	AC
XX	ABL58139;
XX	
XX	30-JUL-2002 (first entry)
XX	
XX	Human phosphorylcholine/ethanolamine transferase 15.29 coding sequence.
DE	
XX	
KW	Human, phosphorylcholine/ethanolamine transferase 15.29; enzyme;
KW	phosphorylcholine; ethanolamine; transferase; tumour; inflammation;
KW	cytostatic; anti-inflammatory; gene; gene therapy; ss.
XX	
OS	Homo sapiens.
XX	
XX	
PH	Location/Qualifiers
Key	411..830
CDS	
FT	/*tag= a
FT	/product= "Human phosphorylcholine/ethanolamine
FT	transferase 15.29"
XX	
XX	
PN	CN1331315-A.

[illegible]

RESULT 41	
ADM03261	
ID	ADM03261 standard; cDNA; 2534 BP.
XX	
AC	ADM03261;
XX	
DT	20-MAY-2004 (first entry)
XX	
DE	Human cDNA of the invention SEQ ID NO:1946.
XX	
KW	ss; gene; human; gene therapy; diagnostic marker; pharmaceutical.
XX	
OS	Homo sapiens.
XX	
PN	EP1347046-A1.
XX	
PD	24-SEP-2003.
XX	
PF	12-APR-2002; 2002EP-00008400.
XX	
PR	22-MAR-2002; 2002JP-00137785.
XX	
PA	(REAS-) RES ASSOC BIOTECHNOLOGY

Qy	660	CATGTGCGGGNAGGAGGAGCCTCAGCCATGCTCACCGTCAGGCGCCATGCTTCCAAGTT	719
Db	1393	TGAGGCGGGNAACTGGTACAAAGGCCACGCTCATTATGACCGCATCCCCATCAAGTT	1452
Qy	720	CATAGAGGCTGTGAGGAATGAAGAGGCCACAGAAAGGGGACACGCCACGCTGGTGTGTA	779
Db	1453	TGTGAGCAACCTCAAAATGTACTGTTGAAGAGAGGAGTCCGCATCGCTGGAGTGTGA	1512
Qy	780	GCTGAAA--GCNAGCGCGCACCGGTGGAGTGGAGGAAGGGGCATCAGAGACCTCAGAGATGG	836
Db	1513	GCTGACATCCNAGGATGTGACATCCGCTGGNAGAGGATGGGCAGCTGCTGATGCATGG	1572
Qy	837	GGACAGACACAGCCTTGAGGCGAGHACGGGTCCAGTGTGAGCTGCAGATCCGTTGGCCCTGGC	896
Db	1573	CATTAGTACAGCATGAACCATGAGGGCAAGCAAGCAGAGCTGATCATCGAGGATGCACA	1632
Qy	897	TGTGGTGGATCCGGGGAGTACTCGTTGTGTGCGGGCAGGA	938
Db	1633	GCTCAGTGATGGTGGCGAGTACACTGTGGTGGCCATGCAGA	1674

```

RESULT 14
ABN40768
ID ABN40768 standard; DNA; 60 BP.
XX
AC ABN40768;
XX
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:13516.
XX
KW Human; mouse; rat; splice transcript; detection; RNA transcript;
KW splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
XX 20-JUL-2001; 2001WO-1B001903.
XX
XX
PR 28-JUL-2000; 2000US-0221607P.
PR 02-MAY-2001; 2001US-0287724P.
XX
XX (COMP-) COMPUGEN INC.
XX
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI: 2002-257383/30.

```

CC - and pathology-specific genes such as those genes only expressed in
CC specific tissue under a specific pathological condition; to detect
CC developmental specific genes; and to detect RNA transcripts and splice
CC variants of a transcriptome of a patient suffering from a particular
CC disorder. ABN27253 to ABN59589 represent oligonucleotide sequences from
CC rats, humans and mice, which are used in the exemplification of the
CC present invention. N.B. The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 60 BP; 18 A; 13 C; 20 G; 9 T; 0 U; 0 Other;

Query Match 6.0%; Score 60; DB 6; Length 60;
Best Local Similarity 100.0%; Pred. No. 7e-05;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 703 GCCATGCGCTTCCAAAGTTCATAGAGGGTCTGAGGAATGAAGAGGCCACAGAGGGGACACG 762
|||
Db 1 GCCATGCGCTTCCAAAGTTCATAGAGGGTCTGAGGAATGAAGAGGCCACAGAGGGGACACG 60

RESULT 15
ABX13540
ID ABX13540 standard; DNA; 93801 BP.
XX
AC ABX13540;
XX
DT 05-JUN-2003 (first entry)
XX
DE Human RGS11 DNA.
XX
KW RGS11; human; screening; cardiant; antianginal; gene therapy; gene;
KW heart disorder; cardiac ischaemia; heart failure; angina; de.
XX
OS Homo sapiens.
XX
PN W02002103355-A1.
XX
PD 27-DEC-2002.
XX
PF 17-JUN-2002; 2002WO-JP006019.
XX
PR 18-JUN-2001; 2001JP-00183038.
XX
PA (TAKE) TAKEDA CHEM IND LTD.
XX
PI Koyama N, Tanida S, Yamamoto K;
XX
DR WPI; 2003-167557/16.
XX
DR P-PSDB; ABG74786.
PT Screening compounds regulating RGS11 expression and activity for
PT prevention and treatment of heart disease.
XX
PS Example 1; Page 262-316; 321pp; Japanese.
XX
CC This invention describes a novel method for screening compounds for their
CC ability to regulate the activity and expression of human RGS11 and its
CC partial peptides and salts, by observing the expression or activity of
CC RGS11 in the presence or absence of the test compound. The products of
CC the invention have cardiant and antianginal activity and can be used for
CC gene therapy. The methods and compositions are useful in the prevention,
CC treatment and diagnosis of heart disorders such as cardiac ischaemia,
CC heart failure and angina. This sequence encodes the human RGS11 protein
CC described in the disclosure of the invention
XX
SQ Sequence 93801 BP; 30092 A; 19478 C; 21592 G; 22639 T; 0 U; 0 Other;

Query Match 5.9%; Score 59; DB 9; Length 93801;
Best Local Similarity 43.4%; Pred. No. 0.00067;
Matches 376; Conservative 0; Mismatches 485; Indels 6; Gaps 2;

Qy 141 GCAGGAGAGACCTCAGGCCACGCTCACCATCAGGGGTCTGCCCTGCCAGGTTTCATAGAAGA 200

Db 31821 GGAGCTGAGCTCACAGCACAGCTGAAGGCTTTAGAGGCCGATCCCTACTTCACTGTGAA 31880
QY 201 TGTGAAAAACAGGAGGCCAGAGAAAGGGCCACGCGTGTGCTGCAGTGTGTGAGCTGAACAG 260
Db 31881 ATTACATGACAAACTGCGTGCAGAGGATGAGATTACTTTGAAGTGTGAAGTGAGCAA 31940
QY 261 TGCAGCCCTGTGGAGTGGAGAAAGGGTCTGAGACCCTCAGAGATGGGGACAGATACAG 320
Db 31941 AGATGTACCAGTGAATGTTCAAGATGGTGAAGAGATTGTCCTTCAACCAATATTC 32000
QY 321 CCTGAGGCAGGAGGAGCTAAATGTGAGCTGCGATTCGTGGCCTGGCCATGGCAGACAC 380
Db 32001 TATCAAGGCAGATGCGCTCGCGCATCTTAAATAATCAAAAAGCGGACCTTAAAGATAA 32060
QY 381 TGGGAGTACTCGTGTGGGCGAGGAGGACCTCGGCTATGCTCAGCTCAGGGC 440
Db 32061 AGGCGAATATGTGTGACTGTGGCACAGACAGACCAAGGCAATGTACTGTTGAGGC 32120
QY 441 TCTACCCATCAAGTTTACAGAGGGTCTGAGGAACGAGAGGCCACAGAAAGGGGCAACAGC 500
Db 32121 TCGACTAATAAAGTGAAGAACCTCTGTACGAGTAGAGGTGTTTGTGGTGAACAGC 32180
QY 501 CGTGTGCGGTGTGAGCTGAGCAAGATGGCCCCCGTGGAGTGGTGAAGGGGCATG---A 557
Db 32181 CCACTTTGAATTTGAACTTTCTGAACCTGATGTTTCCAGGCCAGTGGAAAGCTGAAAGGACA 32240
QY 558 GACCCCTCAGAGATGGAGACAGACACAGCCTGAGGCGAGGCGGGCCAGGTGTGAGCTGCA 617
Db 32241 GCCTTTGACAGCTTCCCTGACTGTGAATCATTTGAGGATGGAAAGAGCATNTTCTGAT 32300
QY 618 GATCCGCGCCTGTGGCAGAGGACGCTGGGAGTACCTGTGCATGTGCGGGAAAGGAGAG 677
Db 32301 CCTTCATAACTGTCACTGCTGGGTATGACAGGAGAGGTTTCTTCCAGGCTGCTAATGCCAA 32360
QY 678 GACCTCAGCCATGCTCACCGTCAGGCCATGCCCTCCAAGTTTCATAGAGGCTCTGAGGAA 737
Db 32361 ATCTGACGCCAATCTGAAGTGAAGAAATTGCCTCTTATCTTCATCACACCTCTCAGTGA 32420
QY 738 TGAAGAGGCCACAGAAAGGGGACACGGCCACGCTGTGGTGTGAGCTGAGCAAGGGCGC--- 794
Db 32421 TGTAAAGTCTTCGAGAAAGATGAGGCTAAGTTTGAGTGTGAAGTATCCAGGGAGCCCAA 32480
QY 795 ACCGGTGGAGTGGAGGAAGGGGCATGAGACCCCTCAGAGATGGGACAGACACAGCCTGAG 854
Db 32481 AACATTCCGTTGGCTTAAAGGAAACCCAGGAAATCACAGGTGATGACAGATTGAGCTTAT 32540
QY 855 GCAGACGGGTCCAGGTGTGAGCTGCAGATCCGTGGCTGCTGTGGTGGATCCGGGGA 914
Db 32541 AAAGATGGCACTAAGCAATTCAATGGTGATCAAGTCACTGCTCTTTTGAAGATGAAGCAA 32600
QY 915 GTACTCGTGTGTGTCGGGCGAGGAGGACCTCAGCCACACTCACTGTGAGGGCCCTGCC 974
Db 32601 ATACATGTTTGAAGCTGAAGATAAGCACAAAGTGGCAAACTGATCATTTGAAGAAATCCG 32660
QY 975 TGCAGATTTCATAGAAGATGTGAATAA 1001
Db 32661 GCTCAAAATTCCTACCCCTCTCAAGA 32687

Search completed: March 20, 2005, 18:54:52
Job time : 585.044 secs

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	248.4	24.8	2768	4	US-09-774-528-426	Sequence 426, Appl
2	69.6	7.0	7218	1	US-08-232-463-14	Sequence 14, Appl
C 3	55.4	5.5	3766	4	US-09-949-016-1651	Sequence 1651, Ap
4	55.4	5.5	3778	4	US-09-949-016-359	Sequence 359, App
C 5	52.6	5.3	1926	3	US-09-249-585A-4	Sequence 4, Appli
C 6	52.6	5.3	1931	2	US-09-130-114-2	Sequence 2, Appli
7	49.6	5.0	5207	4	US-09-858-664A-1	Sequence 1, Appli
8	49.6	5.0	5207	4	US-10-274-978-1	Sequence 1, Appli
9	49.6	5.0	5207	4	US-10-274-978-3	Sequence 3, Appli
10	49.6	5.0	5207	4	US-10-697-263-1	Sequence 1, Appli
11	49.6	5.0	5207	4	US-10-697-263-3	Sequence 3, Appli
12	49.2	4.9	1505	1	US-07-915-246-1	Sequence 1, Appli
13	47.2	4.7	3575	4	US-09-023-655-1395	Sequence 1395, Ap
14	47	4.7	18955	4	US-09-949-016-13343	Sequence 13343, A
15	47	4.7	30678	4	US-09-949-016-13818	Sequence 12818, A
16	46.8	4.7	1926	3	US-09-249-585A-2	Sequence 2, Appli
17	46.8	4.7	1926	4	US-09-410-399-3	Sequence 3, Appli
18	46.8	4.7	2580	3	US-09-050-863-2	Sequence 2, Appli
19	46.8	4.7	2580	3	US-09-359-081-2	Sequence 2, Appli
C 20	46.8	4.7	5452	2	US-09-130-114-1	Sequence 1, Appli
C 21	46.8	4.7	8705	3	US-09-647-344A-14	Sequence 14, Appl
C 22	46.8	4.7	9600	3	US-08-910-647-1	Sequence 1, Appli
23	46.8	4.7	9600	3	US-09-620-925-1	Sequence 1, Appli
24	46.8	4.7	10596	1	US-07-884-811-15	Sequence 15, Appl
25	46.8	4.7	10596	1	US-07-884-811-15	Sequence 15, Appl
26	46.8	4.7	10596	1	US-08-087-783A-15	Sequence 15, Appl
27	46.8	4.7	10596	1	US-08-194-088B-15	Sequence 15, Appl

QY 249 CCAGGCCACAGACGGAGGTGACGTGGTGTACAAAGATGGGAAGAGCTGAGCTCCAGGCTCAA 308
Db 2054 AGATCAGGCTGCGGCTGTGACGTGGCTGAAGATGTCACACACTGTCCCGAGGCCCA 2113
QY 309 AGTGGGCATGAGAGTCAAAAGGTGCAACAGAGGTGGTGTGTCACAGCGCGGCAAGC 368
Db 2114 GTATGAGGTGCAGGCATCGCGCGGGCGGGTGTCTTGTGCGAGATGTGGCCCGGA 2173
QY 369 AGATGCTGGGAGTACAGCTGTGAGGCTGGGGCCAGAGAGTCTCTTCCACTGCACAT 428
Db 2174 CGATGACGCCCTCTACGATGGCTGACCGCGGGGCCGATCGCTACAGCTTCCGT 2233
QY 429 CACAGAGCCCAAGGGGTGTTTGCAGAGGAGCAGTCAAGTGCATAATGAGTGCAGGCTGA 488
Db 2234 GCAAGGCCCTCGCGCTTCTTCGACACAGGACATGGCGGCGCAGCTGT---GTGATGCCGT 2290
QY 489 GCGGGGACCATGCTCATCTGAGCTGTGAGTGGGCCAGCCCAAGCGAGGTGACGTG 548
Db 2291 GCGTGGGGCCCGGCGCAGTGTGAGTGTGAGACCTCCGAAGCCACGTCACGTCGACTG 2350
QY 549 GTACAAGGACGGAAGAGCTGAGCTCCAGCTCAAAAGTACGATGGAGGTCAAGGGCTG 608
Db 2351 GTACAAGATGCATGAGCTGGGCACTCCGCTGAGCGCTTCTTCAGAGGAGATGTGG 2410
QY 609 CACAGC---AAGCTGTGATGCAGCAGGTGGGCAAGCAGATGCTGGGAGTACAGCTG 665
Db 2411 GACGGGACCGCTGGTGGCAGCCACAGTCCACAGGCGAGGATGAAGCACCTACTCTG 2470
QY 666 CGAGCTGGGGCCAGAGTCTCTTCACTGACATCAAGAGCCCAAGCGAGTGT 725
Db 2471 CCGCTGGCGGAGGACTCTGTGGACTTCGGCTCCGCTCTGTGAGCCCAAGCGGTGT 2530
QY 726 TGCCAAAGAGCAGTGTGGTGCATAATGAGTGGGACTGAGGCGAGGCGCAGTCCACACT 785
Db 2531 TGCCAAAGAGCAGCGCGGTGAGGAGGTGAGCTGAGGTGGGGCCAGCGCACGCT 2590
QY 786 GAGCTGTGAGGTGGCCCGCCAGCCAGACAGAGGTGACGTGGTACAAAGATGGGAAGCT 845
Db 2591 GAGCTGTGAGGTGGCCCGCCAGCCAGATGAGGTGACGTGGTACAAAGCGGGAAGT 2650
QY 846 GAGCTCCAGTTCGAAGTGGCATAGAGCTCGGGCTGCATGGCGGAGCTGGTGTGCA 905
Db 2651 GAGCTCCAGCTCGAAGTGCAGTGGAGGCTGTGGGCTGTATGCGAGGCTGGTGTGCA 2710
QY 906 CGAGCAGCGCCAGGAGATGCTGGGGAGTACACCTGTGAGGCTGGGG 953
Db 2711 GCAGGTGGCCAGGCAGACTCCGGATAGTACAGCTGTGAAGCCAGGTG 2758

RESULT 2

US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367

GENERAL INFORMATION:

; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/232,463

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/07/935,313

; FILING DATE:

; APPLICATION NUMBER: EP 91 114 300.6

; FILING DATE: 26-AUG-1991

; ATTORNEY/AGENT INFORMATION:

; NAME: BENT, Stephen A.

; REGISTRATION NUMBER: 29,768

; REFERENCE/DOCKET NUMBER: 30472/114 IMMU

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (703) 836-9300

; TELEFAX: (703) 683-4109

; TELEX: 899149

; INFORMATION FOR SEQ ID NO: 14:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7218 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; CLONE: pTZgpt-Fls

; US-08-232-463-14

Query Match 7.0%; Score 69.6; DB 1; Length 7218;
Best Local Similarity 4.5%; Pred.No. 5.2e-08;

Matches 18; Conservative 233; Mismatches 147; Indels 0; Gaps 0;

QY 28 TTGAAGTGCATGTAGAGGCCAAGGCTGCAGCGAGGCTGTGTGTGCAGCAGGAGGC 87
Db 1441 TTGTACRR 1382

QY 88 AAGAGGATCGCGGGGACTACAGCTGCAGCGAGGCGGCGCAGAGGCTCTCTCCGCGTG 147
Db 1381 RRR 1322

QY 148 CACATCACAGACGCCAAGATGATGTTTGCAGAGGAGCAGTCAGTGCAATAGAGTGCAG 207
Db 1321 RRR 1262

QY 208 GCTGAGCGGGGCGCATGCTGCATGTGAGTGTGAGTGCCGCCAGCCAGCAGGAGTG 267
Db 1261 RRR 1202

QY 268 ACCTGGTACAGGATGGGAAGAGCTGAGCTCCAGCTCAAAAGTGGCATGGAGGTCAA 327
Db 1201 RRR 1142

QY 328 GGGTGACACGAGGCTGTGTGCTGCCACAGCGGCAAGCAGATGCTGGGAGTACAGC 387
Db 1141 RRR 1082

QY 388 TGTGAGCTGGGGCCAGAGAGTCTCTTCCACTGCA 425
Db 1081 RRR 1044

RESULT 3

US-09-949-016-1651

; Sequence 1651, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768


```
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1651
; LENGTH: 3766
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1651

Query Match          5.5%; Score 55.4; DB 4; Length 3766;
Best Local Similarity 48.8%; Pred. No. 0.00019;
Matches 186; Conservative 0; Mismatches 186; Indels 9; Gaps 1;

Qy 481 CAGGCTGAGGCGGGACCACTGCCATGCTGAGCTGTGAGTGGCCCGCCAGACGAG 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 868 CAGGTTGACAAAGAGGAGGAGAGTGGTTGTTGAGAGCTGGCAGATCCAAAGTTGGAG 927
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 541 GTGACGTGGTACAAAGGACGGGAAGAGCTGAGCTCCAGCTCAAAAGTACGATGGAGGTC 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 928 GTGAAATGGTATAAAATGGTCAAGAAATTCGACCCAGTACCAATACATCTTTGAACAC 987
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 601 AAGGCTGCACACAAAGGCTGGTAGTCAGAGGTCGGGCAAGAGATGCTGGGGAGTAC 660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 988 AAAGGATGCCAGAGATCTCTTTTATCAATACTGTACAGATCAGAGATGATTCAGAGTAT 1047
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 661 AGCTGCGAGGCTGGGGCCAGAGAGTCTCTTTTCAACTGCACATCACAGAGCCCAAGCA 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1048 TATGTGACAGCCGGTGATGAGAAATGTTCCACTGAGCTCTTCGTAAGAGAGCCTCCAATT 1107
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 721 GTGTTTGCACAGGAGCAGTGGTGCATATAGAGTGGCGACTGAGGCGAGGGCCAGTGCC 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1108 ATGCTGACCAACAGCTGGAAGATACAACCTG-----CTTATTGGGGAGAGACTG 1158
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 781 ACACGTAGCTGTGAGGTGGCCAGGCCAGACAGAGTGAGCTGTACAAAGATGGGAAG 840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1159 GAATTAGAAATGTAGGTGTCTGAAAGATGATGCCAAATGTAAATGTTTAAAGATGGTGA 1218
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 841 AAGCTGAGCTCCAGTTCGAAA 861
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1219 GAGATTATCTCTGGTCCAAAA 1239
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 5
US-09-249-585A-4/c
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match          5.3%; Score 52.6; DB 3; Length 1926;
Best Local Similarity 41.9%; Pred. No. 0.00078;
Matches 316; Conservative 0; Mismatches 439; Indels 0; Gaps 0;

Qy 200 AGGTGACAGCTGAGCGGGGCCAGTGCCATGCTGAGCTGTGAGTGCCCGCCAGACG 259
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1125 AGGTGCTGGAGAGGGGACCGAGAAAGTGTGAAAGGGGGGACCGAGAAAGTGCAGGAGA 1066
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 260 CGGAGGTGACGTGTACAAAGGATGGGAAGAGCTGAGCTCCAGCTCAAAAGTGGGCATGG 319
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1065 TGGGGCCCGGAGGTGATGGAGAGCTGGGGCCGAGGTGATGGAGAGCTGGGGCCGG 1006
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 320 AGGTCAAAAGGTCACACGGAGGCTGGTGTGCTGCCACAGGCGGGCAAGAGATGCTGGGG 379
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1005 AGGTGACGGAGGACTGGGGCCGAGGTGGAGGACGAGGACGGGGAGCAGGACGGGG 946
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 380 AGTACAGCTGTGAGGCTGGGGGCCAGAGAGTCTCTTCCACTGTCACATCACAGAGCCCA 439
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 945 AGGAGGACGAGGACGGGGAGGACCGGGGAGGACGAGGACGGGGAGGACCGGGAGGACGAGG 886
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 440 AGGGGGTGTTCGCAAGGAGCAGTCAGTGCATATAGGTGTCAGGCTGAGGCGGGGACCA 499
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1651
; LENGTH: 3766
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-1651

Query Match          5.5%; Score 55.4; DB 4; Length 3766;
Best Local Similarity 48.8%; Pred. No. 0.00019;
Matches 186; Conservative 0; Mismatches 186; Indels 9; Gaps 1;

Qy 481 CAGGCTGAGGCGGGACCACTGCCATGCTGAGCTGTGAGTGGCCCGCCAGACGAG 540
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 868 CAGGTTGACAAAGAGGAGGAGAGTGGTTGTTGAGAGCTGGCAGATCCAAAGTTGGAG 927
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 541 GTGACGTGGTACAAAGGACGGGAAGAGCTGAGCTCCAGCTCAAAAGTACGATGGAGGTC 600
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 928 GTGAAATGGTATAAAATGGTCAAGAAATTCGACCCAGTACCAATACATCTTTGAACAC 987
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 601 AAGGCTGCACACAAAGGCTGGTAGTCAGAGGTCGGGCAAGAGATGCTGGGGAGTAC 660
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 988 AAAGGATGCCAGAGATCTCTTTTATCAATACTGTACAGATCAGAGATGATTCAGAGTAT 1047
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 661 AGCTGCGAGGCTGGGGCCAGAGAGTCTCTTTTCAACTGCACATCACAGAGCCCAAGCA 720
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1048 TATGTGACAGCCGGTGATGAGAAATGTTCCACTGAGCTCTTCGTAAGAGAGCCTCCAATT 1107
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 721 GTGTTTGCACAGGAGCAGTGGTGCATATAGAGTGGCGACTGAGGCGAGGGCCAGTGCC 780
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1108 ATGCTGACCAACAGCTGGAAGATACAACCTG-----CTTATTGGGGAGAGACTG 1158
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 781 ACACGTAGCTGTGAGGTGGCCAGGCCAGACAGAGTGAGCTGTGTAAGATGGGAAG 840
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1159 GAATTAGAAATGTAGGTGTCTGAAAGATGATGCCAAATGTAAATGTTTAAAGATGGTGA 1218
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 841 AAGCTGAGCTCCAGTTCGAAA 861
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1219 GAGATTATCTCTGGTCCAAAA 1239
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
US-09-949-016-359
; Sequence 359, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 359
; LENGTH: 3778
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-359

Query Match          5.5%; Score 55.4; DB 4; Length 3778;
Best Local Similarity 48.8%; Pred. No. 0.00019;
Matches 186; Conservative 0; Mismatches 186; Indels 9; Gaps 1;
```



```
QY 640 AAAGCAGATGCTGGGAGTACAGCTGCGAGGC 671
Db 460 TCGAAGGATGCCGCGCTTTACACCTGCTGCG 491

RESULT 8
US-10-274-978-1
; Sequence 1, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,978
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Human
US-10-274-978-1

Query Match 5.0%; Score 49.6; DB 4; Length 5207;
Best Local Similarity 57.9%; Pred. No. 0.0065;
Matches 88; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 520 GTGCCCCAGCCAGCAGGAGTACGCTGTTACAAAGGACGGGAAGCTGAGCTCCAGC 579
Db 340 GAGGCGACCCACACAGCCCTCGTGACCTGGTACAAAGGACAGGCTCCAGCTGGTGACAGC 399

QY 580 TCAAAAGTACGATGAGGTCAAGGCTGCACACGAGGCTGTAGTGCAGCAGGTGGGC 639
Db 400 ACCCGGCTTAGCCAGCAGCAAGAGGACCACTACTCCCTGGTGTCTGAGGCATGTGGCC 459

QY 640 AAAGCAGATGCTGGGAGTACAGCTGCGAGGC 671
Db 460 TCGAAGGATGCCGCGCTTTACACCTGCTGCG 491

RESULT 9
US-10-274-978-3
; Sequence 3, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,978
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Human
US-10-274-978-3

Query Match 5.0%; Score 49.6; DB 4; Length 5207;
Best Local Similarity 57.9%; Pred. No. 0.0065;
Matches 88; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 520 GTGCCCCAGCCAGCAGGAGTACGCTGTTACAAAGGACGGGAAGCTGAGCTCCAGC 579
Db 340 GAGGCGACCCACACAGCCCTCGTGACCTGGTACAAAGGACAGGCTCCAGCTGGTGACAGC 399

QY 580 TCAAAAGTACGATGAGGTCAAGGCTGCACACGAGGCTGTAGTGCAGCAGGTGGGC 639
Db 400 ACCCGGCTTAGCCAGCAGCAAGAGGACCACTACTCCCTGGTGTCTGAGGCATGTGGCC 459

QY 640 AAAGCAGATGCTGGGAGTACAGCTGCGAGGC 671
Db 460 TCGAAGGATGCCGCGCTTTACACCTGCTGCG 491

RESULT 10
US-10-697-263-1
; Sequence 1, Application US/10697263
; Patent No. 6812014
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,263
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-697-263-1

Query Match 5.0%; Score 49.6; DB 4; Length 5207;
Best Local Similarity 57.9%; Pred. No. 0.0065;
Matches 88; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 520 GTGCCCCAGCCAGCAGGAGTACGCTGTTACAAAGGACGGGAAGCTGAGCTCCAGC 579
Db 340 GAGGCGACCCACACAGCCCTCGTGACCTGGTACAAAGGACAGGCTCCAGCTGGTGACAGC 399

QY 580 TCAAAAGTACGATGAGGTCAAGGCTGCACACGAGGCTGTAGTGCAGCAGGTGGGC 639
Db 400 ACCCGGCTTAGCCAGCAGCAAGAGGACCACTACTCCCTGGTGTCTGAGGCATGTGGCC 459

QY 640 AAAGCAGATGCTGGGAGTACAGCTGCGAGGC 671
Db 460 TCGAAGGATGCCGCGCTTTACACCTGCTGCG 491

RESULT 11
US-10-697-263-3
; Sequence 3, Application US/10697263
; Patent No. 6812014
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,263
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
```



```
/ PRIOR FILING DATE: 2001-05-17
/ PRIOR APPLICATION NUMBER: 09/711,134
/ PRIOR FILING DATE: 2000-11-14
/ NUMBER OF SEQ ID NOS: 34
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 3
/ LENGTH: 5207
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-697-263-3

Query Match      5.0%; Score 49.6; DB 4; Length 5207;
Best Local Similarity 57.9%; Pred. No. 0.0065;
Matches 88; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 520 GTGGCCACCCAGCAGGAGGTGACGTGATACAGGAGCGGAAAGAGTGAAGTCCAGC 579
Db 340 GAGGCGGACCCACAGCCCTCGGTGACCTGGTACAGGAGCAGCGTCCAGCTGGTGGACAGC 399
QY 580 TCAAAAGTACGATGAGGTCAAGGCTGACACGAGGCTGATAGTGCAGCAGGTGGGC 639
Db 400 ACCGCGTITAGCAGCAGCAGCAAGAGGACCAATACCTCCCTGGTCTGAGGATGTGGCC 459
QY 640 AAGCAGATCTGGGAGTACAGCTGCGAGGC 671
Db 460 TCGAAGGATGCCGCGGTTTACACCTGCTGCTGGC 491

RESULT 12
US-07-915-246-1
/ Sequence 1, Application US/07915246
/ Patent No. 5401836
/ GENERAL INFORMATION:
/ APPLICANT: Baezzyński, Chris L.
/ APPLICANT: Fallis, Lynne
/ APPLICANT: Bellmare, Guy
/ APPLICANT: Boivin, Rodolphe
/ TITLE OF INVENTION: A BRASSICA REGULATORY SEQUENCE FOR
/ ROOT-SPECIFIC OR ROOT-ABUNDANT GENE EXPRESSION
/ NUMBER OF SEQUENCES: 5
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: McAndrews, Held, and Malloy
/ STREET: 500 W. Madison St. 34th Floor
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60661
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/915,246
/ FILING DATE: 19920716
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Pochopien, Donald J.
/ REGISTRATION NUMBER: 32,167
/ REFERENCE/DOCKET NUMBER: 91 P 1125
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312 707-8889
/ TELEFAX: 312 707-9155
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1505 base pairs
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: double
/ TOPOLOGY: linear
/ MOLECULE TYPE: DNA (genomic)
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ ORIGINAL SOURCE:

/ ORGANISM: Brassica napus
/ STRAIN: Westar
/ DEVELOPMENTAL STAGE: Somatic
/ TISSUE TYPE: Root
US-07-915-246-1

Query Match      4.9%; Score 49.2; DB 1; Length 1505;
Best Local Similarity 45.0%; Pred. No. 0.0054;
Matches 226; Conservative 0; Mismatches 273; Indels 3; Gaps 1;

QY 69 GGTGCTGACGACGAGGCAAGACGATGCGGGGACTACAGCTGCGAGGCGCAGGGGCCA 128
Db 876 GGTGCTGAGCTGATACGAGGCGGAGGAGCTGGAGGACATGCTGGAGGTGGAGGAGC 935
QY 129 GAGGGTCTCTTCGCGCTGCACATCACAGAGCCCAAGATGATGTTTCCAAAGGAGCAGTC 188
Db 936 GGAGGAGGTGGTGGTGGAGGAGCTGGCGGTGCGCATGCTGGTGGATACGGTGGTGGAGAA 995
QY 189 AGTCATAATGAGGTGAGGCTGAGGCGGGGCCAGTGCCATGCTGAGCTGTGAGGTGGC 248
Db 996 GGTGCTGGTCTGAGGAGGATATGAGGTGGCGGTGCGGTGCGATGGAGGTGGTGGGA 1055
QY 249 CCAGGCCACGACGAGGTGACGTGGTACAAGGATGGAAGAGTGAAGTGAAGTCCAGCTCAA 308
Db 1056 GCGCGTGGAAATGAGGCGGTGGAGGAGGTGGAGTGCACACGCTGGTGATA---CGGT 1112
QY 309 AGTGGGATGAGGTCAAGGGTGCACACGAGGCTGGTGTGCCACAGGCGGGGCAAGC 368
Db 1113 GGTGGAAGAGCGCTGGTGTGAGGAGGATATGAGGTGGCGGTGCGGTGCGATGGGA 1172
QY 369 AGATGCTGGGAGTACAGCTGTGAGGCTGGGGGCCAGAGAGTCTCTTCCACCTGCACAT 428
Db 1173 GGTGCTGAGCGGGGAAAAGAGGCGGTGGAGGAGGAGTCTTGGCGCGGTGGAGCT 1232
QY 429 CACAGAGCCCAAGGGGTGTTTTCGAGAGGAGCAGTCAATGATGATGAGGTGAGGCTGA 488
Db 1233 CACGCTGGTGTATGTTGTCGCGGAGGTGGAGTGCAGAGGATACGCTGGTGGTGGGA 1292
QY 489 GCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCAGCCGCCACAGCGAGGTGAGTG 548
Db 1293 GAAGGAGGACCGTGTGAGGAGGCGGTGGTGGAGCTGGAGTGGCGGAGGAGGA 1352
QY 549 GTAAAGGACGGGAAGAAGCTG 570
Db 1353 GCGGAGGTTTATGAGCTGCTG 1374

RESULT 13
US-09-023-655-1395
/ Sequence 1395, Application US/09023655
/ Patent No. 6607879
/ GENERAL INFORMATION:
/ APPLICANT: Cocks, Benjamin G.
/ APPLICANT: Susan G. Stuart
/ APPLICANT: Jeffrey J. Sellhamer
/ TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
/ EXPRESSION
/ NUMBER OF SEQUENCES: 1508
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
/ STREET: 3174 PORTER DRIVE
/ CITY: PALO ALTO
/ STATE: CALIFORNIA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/023,655
/ FILING DATE: HERewith
```


Qy	239	GTGAGGTGGCCACGAGCCCGACGAGGTGACGTGTGTTACAAGGATGGGAAAGACTGAGCT	298
Db	15150	AGGAGGAGGGGAGGAGGAGGGGAGGAGTAAAGAGTGAAGGGAGGCGAAGGGGAGGA	15209
Qy	299	CCAGCTCAAAAGTGGGCTGAGGTCAAAAGGTTGCACACGGAGGCTGTGTGCTGCCACAGG	358
Db	15210	GTAGGAGGAGGAGGAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	15269
Qy	359	CGGGCAAGCAGATGCTGGGGAGTACAGCTGTGAGGCTGGGGGCCAGAGAGTCTCCTTCC	418
Db	15270	AGGGNAGGAGGAGGAGGAGGGGAGAGAGAGGGAGGAGTAGGAGGGAGCAGAGGGGG	15329
Qy	419	ACCTGCACATCACAGAGCCCAAGGGGGTGTTCGGAAGGAGCAGTCAGTGCATAATGAGG	478
Db	15330	AGGAATAGGAGGGGAGGAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	15389
Qy	479	TGCAGGCTGAGGGGGGACCACTGCTCATGCTGAGCTGTGAGTGGCCAGCCCCACAGCG	538
Db	15390	GGGAGGAGTAGGAGGTGAGCA-GAAGGGGAGGAGTAGGAGGGGAGGAGGAGGGGAGGG	15448
Qy	539	AGGTGACGTGTGTACAGGACGGGAAGAGCTGAGCTCCAGCTCAAAAGTACGCATGGAGG	598
Db	15449	AGAGGAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT	15508
Qy	599	TCAAGGGCTGCACACGAAGGCTGGTAGTGCACAGGTGGGCAAAAGCAGATGCTGGGGAGT	658
Db	15509	AGGAGGGGAGCAGAAAGGGGAGGAGTAGGAGGGGAGGAGGAGGAGGAGGAGGAGGAGG	15568
Qy	659	ACAGCTGCGAGGCTGGGGGCCAGAGAG	685
Db	15569	GAGGAGAAAGGGGAGGAGTAGGAGGGAG	15595

Search completed: March 21, 2005, 14:38:40
Job time : 179.678 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:32:32 ; Search time 627.052 Seconds
(without alignments)
9503.485 Million cell updates/sec

Title: US-10-077-130-4_COPY_3120_4120
Perfect score: 1001
Sequence: 1 aaagatgggaagagctgag.....tcagagcccaaggcggtgtt 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues
Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications NA:*
- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
 - 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
 - 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
 - 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
 - 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
 - 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
 - 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
 - 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
 - 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
 - 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
 - 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
 - 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
 - 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
 - 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
 - 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
 - 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
 - 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
 - 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
 - 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
 - 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
 - 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
 - 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	23907	13 US-10-077-130-6	Sequence 6, Appli
2	1001	100.0	24120	13 US-10-077-130-4	Sequence 4, Appli
3	965	96.4	20489	18 US-10-723-860-5701	Sequence 5701, Ap
4	248.4	24.8	2768	17 US-10-120-988-426	Sequence 426, App
5	185.6	18.5	578	13 US-10-027-632-272054	Sequence 272054,
6	185.6	18.5	578	17 US-10-027-632-272054	Sequence 22, Appl
7	123.2	12.3	1645	9 US-09-726-643-22	Sequence 22, Appl
8	123.2	12.3	1645	13 US-10-042-141-22	Sequence 22, Appl
9	123.2	12.3	1645	19 US-10-919-272-22	Sequence 22, Appl
10	123.2	12.3	2170	17 US-10-264-049-850	Sequence 850, App
11	97.4	9.7	14061	17 US-10-093-463-73	Sequence 73, Appl

12	97.4	9.7	14109	17	US-10-093-463-71	Sequence 71, Appl
13	95.8	9.6	2534	17	US-10-108-260A-1946	Sequence 1946, Ap
14	95.8	9.6	7564	18	US-10-476-397-15	Sequence 15, Appl
15	88.4	8.8	2254	9	US-09-726-643-44	Sequence 44, Appl
16	88.4	8.8	2254	13	US-10-042-141-44	Sequence 44, Appl
17	88.4	8.8	2254	19	US-10-919-272-44	Sequence 44, Appl
18	68	6.8	716	10	US-09-822-846-117	Sequence 117, App
19	66.8	6.7	9591	10	US-09-764-891-6043	Sequence 6043, Ap
20	66.8	6.7	9591	14	US-10-091-438-268	Sequence 268, App
21	66.8	6.7	12415	10	US-09-764-891-6044	Sequence 6044, Ap
22	66.8	6.7	12415	14	US-10-091-438-269	Sequence 269, App
23	66.8	6.7	20565	10	US-09-764-891-6045	Sequence 6045, App
24	66.8	6.7	20565	14	US-10-091-438-270	Sequence 270, App
c	64.8	6.5	400	17	US-10-242-535A-8373	Sequence 8373, Ap
25	64.8	6.5	400	17	US-10-085-783A-8373	Sequence 8373, Ap
26	64.8	6.5	400	17	US-10-085-783A-8373	Sequence 1, Appl
27	62	6.2	81940	9	US-09-759-508B-1	Sequence 1052, Ap
28	62	6.2	81940	10	US-09-960-706-1092	Sequence 724, App
29	62	6.2	81940	10	US-09-873-319-724	Sequence 132, App
30	62	6.2	81940	18	US-10-723-860-132	Sequence 1, Appl
31	62	6.2	81940	18	US-10-656-873A-1	Sequence 1034, Ap
32	62	6.2	82027	19	US-10-278-698-1034	Sequence 1034, Ap
33	62	6.2	82027	19	US-10-278-698-1045	Sequence 116, App
34	61.6	6.2	1266	10	US-09-822-846-116	Sequence 1001, Ap
c	61.4	6.1	2884	17	US-10-094-749-1001	Sequence 2224, Ap
35	55.4	5.5	3778	18	US-10-723-860-2224	Sequence 21853, A
36	55.4	5.5	3966	18	US-10-357-930-21853	Sequence 25171, A
37	55.4	5.5	3966	18	US-10-357-930-25171	Sequence 27704, A
38	55.4	5.5	3966	18	US-10-357-930-27704	Sequence 6456, Ap
39	55.4	5.5	3966	18	US-10-357-930-27704	Sequence 11, Appl
40	54.2	5.4	2029	18	US-10-723-860-6456	Sequence 13, Appl
41	53.8	5.4	3535	17	US-10-287-092-11	Sequence 4453, Ap
42	53.8	5.4	3778	17	US-10-287-092-11	Sequence 8268, Ap
43	52.6	5.3	3613	18	US-10-723-860-4453	Sequence 534, App
44	52.6	5.3	4176	18	US-10-723-860-8268	
45	49.6	5.0	3935	17	US-10-108-260A-534	

ALIGNMENTS

RESULT 1

US-10-077-130-6
; Sequence 6, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047P1RCF1(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; TYPE: DNA
; LENGTH: 23907
; ORGANISM: Homo sapiens
US-10-077-130-6

Query Match 100.0%; Score 1001; DB 13; Length 23907;
Best Local Similarity 100.0%; Pred. No. 7.1e-266;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	AAAGATGGGAAGAGCTGAGCTCCAGCTTGAAGCTGATGTAGAGCCAAAGGCTGCAGA	60
Db	3049	AAAGATGGGAAGAGCTGAGCTCCAGCTTGAAGCTGATGTAGAGCCAAAGGCTGCAGA	3108
Qy	61	CGGAGGCTGTGTGTGTCAGCAGCAGGAGCAAGCGGACTACGCTGCAGGCC	120
Db	3109	CGGAGGCTGTGTGTGTCAGCAGCAGGAGCAAGCGGACTACGCTGCAGGCC	3168


```
QY 121 AGGGCCAGAGGGTCTCTTCGCTGCAATCAACAGAGCCCAAGATGATGTTTCAAAG 180
Db |||||
QY 3169 AGGGCCAGAGGGTCTCTTCGCTGCAATCAACAGAGCCCAAGATGATGTTTCAAAG 3228
Db |||||
QY 181 GAGCAGTCAGTCATAATAGAGTGCAGCTGAGGCGGGGGCCAGTGCCTATGCTGAGCTGT 240
Db |||||
QY 3229 GAGCAGTCAGTCATAATAGAGTGCAGCTGAGGCGGGGGCCAGTGCCTATGCTGAGCTGT 3288
QY 241 GAGGTGGCCAGGCGCCAGACGAGAGTGCCTGTTCAAGAGTGGGAAGAGCTGAGCTCC 300
Db |||||
QY 3289 GAGGTGGCCAGGCGCCAGACGAGAGTGCCTGTTCAAGAGTGGGAAGAGCTGAGCTCC 3348
QY 301 AGCTCAAAAGTGGGATGAGAGTCAAAAGGTGCACACGAGAGCTGTGCTGCACAGCG 360
Db |||||
QY 3349 AGCTCAAAAGTGGGATGAGAGTCAAAAGGTGCACACGAGAGCTGTGCTGCACAGCG 3408
QY 361 GGCAAAGCAGATGCTGGGAGTACAGCTGTGAGGCTGGGGGCCAGAGAGTCTCTTCCAC 420
Db |||||
QY 3409 GGCAAAGCAGATGCTGGGAGTACAGCTGTGAGGCTGGGGGCCAGAGAGTCTCTTCCAC 3468
QY 421 CTGCACATCACAGAGCCCAAGGGGCTGTTGGCAAGGAGCAGTCAAGTGCATAATGAGGTG 480
Db |||||
QY 3469 CTGCACATCACAGAGCCCAAGGGGCTGTTGGCAAGGAGCAGTCAAGTGCATAATGAGGTG 3528
QY 481 CAGGCTGAGGGGGGACCACTGCCATGTGAGCTGTGAGGTGGGCCACGCCACAGACGAG 540
Db |||||
QY 3529 CAGGCTGAGGGGGGACCACTGCCATGTGAGCTGTGAGGTGGGCCACGCCACAGACGAG 3588
QY 541 GTGAGCTGTACAAAGCGGAGAGAGTGCAGTCCAGCTCAAAAGTACGATGAGAGTGC 600
Db |||||
QY 3589 GTGAGCTGTACAAAGCGGAGAGAGTGCAGTCCAGCTCAAAAGTACGATGAGAGTGC 3648
QY 601 AAGGCTGCACACGAAGGCTGTGTAGTCAGAGGTGGGCAAGAGCAGATGCTGGGGAGTAC 660
Db |||||
QY 3649 AAGGCTGCACACGAAGGCTGTGTAGTCAGAGGTGGGCAAGAGCAGATGCTGGGGAGTAC 3708
QY 661 AGCTCGAGGCTGGGGCCAGAGAGTCTCTTTCAACTGCACATCACAGAGCCCAAGGCA 720
Db |||||
QY 3709 AGCTCGAGGCTGGGGCCAGAGAGTCTCTTTCAACTGCACATCACAGAGCCCAAGGCA 3768
QY 721 GTGTTTGCACAGGAGCAGTGTGTCATAATGAGGTGGGCAAGAGCAGATGCTGAGGAGGCGGCGGCGG 780
Db |||||
QY 3769 GTGTTTGCACAGGAGCAGTGTGTCATAATGAGGTGGGCAAGAGCAGATGCTGAGGAGGCGGCGGCGGCGG 3828
QY 781 ACATGAGCTGTGAGGTGGGCGCAGAGCAGAGGTGAGCTGTGTCATAAGAGTGGGAAG 840
Db |||||
QY 3829 ACATGAGCTGTGAGGTGGGCGCAGAGCAGAGGTGAGCTGTGTCATAAGAGTGGGAAG 3888
QY 841 AAGCTGAGCTCCAGTTGCAAAAGTGGCATAGAGGTGGGGCTGTCATGCGGCGAGCTGGTG 900
Db |||||
QY 3889 AAGCTGAGCTCCAGTTGCAAAAGTGGCATAGAGGTGGGGCTGTCATGCGGCGAGCTGGTG 3948
QY 901 GTGCAGAGGCGAGGCGCAGAGATGCTGGGAGTACACCTGTGAGGCTGGGGGCCAGCGG 960
Db |||||
QY 3949 GTGCAGAGGCGAGGCGCAGAGATGCTGGGAGTACACCTGTGAGGCTGGGGGCCAGCGG 4008
QY 961 CTCTCTCTTCCACCTGGATGTTTTCAGAGCCCAAGGCGGTGTT 1001
Db |||||
QY 4009 CTCTCTCTTCCACCTGGATGTTTTCAGAGCCCAAGGCGGTGTT 4049
```

RESULT 2

US-10-077-130-4
; Sequence 4, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Accon, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Thereof
; FILE REFERENCE: MPI2001-047PIRCP1(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15

; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 24120
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (1)...(71)
; NAME/KEY: CDS
; LOCATION: (72)...(23978)
; NAME/KEY: 3'UTR
; LOCATION: (23979)...(24120)
US-10-077-130-4

Query Match 100.0%; Score 1001; DB 13; Length 24120;
Best Local Similarity 100.0%; Pred. No. 7,1e-266;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 AAAGATGGGAGAGCTGAGCTCCAGCTTGAAGTGCATGTAGAGGCCAAAGGCTGCAGA 60
Db |||||
QY 3120 AAAGATGGGAGAGCTGAGCTCCAGCTTGAAGTGCATGTAGAGGCCAAAGGCTGCAGA 3179
Db |||||
QY 61 CGGAGGCTGTGTGTGTCAGCAGGCGAGCAAGACGGATGCCGGGACTACAGCTGCGAGGCC 120
Db |||||
QY 3180 CGGAGGCTGTGTGTGTCAGCAGGCGAGCAAGACGGATGCCGGGACTACAGCTGCGAGGCC 3239
QY 121 AGGGGCCAGAGGGTCTCTTCGCTGCACATCAACAGAGCCCAAGATGATGTTTGCAGAG 180
Db |||||
QY 3240 AGGGGCCAGAGGGTCTCTTCGCTGCACATCAACAGAGCCCAAGATGATGTTTGCAGAG 3299
QY 181 GAGCAGTCAGTCATAATGAGTGCAGGCTGAGGCGGGGGCCAGTGCATGCTGAGCTGT 240
Db |||||
QY 3300 GAGCAGTCAGTCATAATGAGTGCAGGCTGAGGCGGGGGCCAGTGCATGCTGAGCTGT 3359
QY 241 GAGGTGGCCAGGCGCCAGAGCGAGGTGACGTGGTACAAAGATGGGAAGAGCTGAGCTCC 300
Db |||||
QY 3360 GAGGTGGCCAGGCGCCAGAGCGAGGTGACGTGGTACAAAGATGGGAAGAGCTGAGCTCC 3419
QY 301 AGCTCAAAAGTGGGATGAGAGTCAAAAGGTGCACACGAGAGGTGTGTGCTGCCACAGCGG 360
Db |||||
QY 3420 AGCTCAAAAGTGGGATGAGAGTCAAAAGGTGCACACGAGAGGTGTGTGCTGCCACAGCGG 3479
QY 361 GGCAAACAGATGCTGGGAGTACAGCTGTGAGCTGGGGGCCAGAGAGTCTCTTCCAC 420
Db |||||
QY 3480 GGCAAACAGATGCTGGGAGTACAGCTGTGAGCTGGGGGCCAGAGAGTCTCTTCCAC 3539
QY 421 CTGCACATCACAGAGCCCAAGGGGCTGTTTTCGAAGGAGCAGTCAAGTGCATAATGAGGTG 480
Db |||||
QY 3540 CTGCACATCACAGAGCCCAAGGGGCTGTTTTCGAAGGAGCAGTCAAGTGCATAATGAGGTG 3599
QY 481 CAGGCTGAGGCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGGCCAGGCCACAGCGAG 540
Db |||||
QY 3600 CAGGCTGAGGCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGGCCAGGCCACAGCGAG 3659
QY 541 GTGAGCTGTGTACAAAGGAGCGGAAGAGCTGAGCTCCAGCTCAAAAGTACGATGAGAGTGC 600
Db |||||
QY 3660 GTGAGCTGTGTACAAAGGAGCGGAAGAGCTGAGCTCCAGCTCAAAAGTACGATGAGAGTGC 3719
QY 601 AAGGGCTGCACACGAAGGCTGTGTAGTGCAGAGGTGGGCAAGAGCAGATGCTGGGGAGTAC 660
Db |||||
QY 3720 AAGGGCTGCACACGAAGGCTGTGTAGTGCAGAGGTGGGCAAGAGCAGATGCTGGGGAGTAC 3779
QY 661 AGCTGAGGCTGGGGGCCAGAGAGTCTCTTTCAACTGCACATCACAGAGCCCAAGGCA 720
Db |||||
QY 3780 AGCTGAGGCTGGGGGCCAGAGAGTCTCTTTCAACTGCACATCACAGAGCCCAAGGCA 3839
QY 721 GTGTTTGCAGAGGAGCAGTGTGTGTCATAATGAGGTGGGAGTGCAGGCGAGGCGGCGGCGGCGG 780
Db |||||
QY 3840 GTGTTTGCAGAGGAGCAGTGTGTGTCATAATGAGGTGGGAGTGCAGGCGAGGCGGCGGCGGCGGCGG 3899
```


Db 3480 GGCAAGCAGATGCTGGGAGGTACAGCTGTGAGGCTGGGGCCAGAGAGTCTCTCTTCCAC 3539
Qy 421 CTGCACATCAGAGCCCAAGGGGTGTTTGCAGAGGAGCAGTCAAGTGCATAAATGAGGTG 480
Db 3540 CTGCACATCAGAGCCCAAGGGGTGTTTGCAGAGGAGCAGTCAAGTGCATAAATGAGGTG 3599
Qy 481 CAGGCTGAGGGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCACGCCCCAGAGGGAG 540
Db 3600 CAGGCTGAGGGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCACGCCCCAGAGGGAG 3659
Qy 541 GTGAGCTGGTACAGGACCGGAAGAGCTGAGCTCCAGCTCAAAAGTAGCAGTGGAGTGC 600
Db 3660 GTGAGCTGGTACAGGACCGGAAGAGCTGAGCTCCAGCTCAAAAGTAGCAGTGGAGTGC 3719
Qy 601 AAGGCTGCACACGAAGGCTGTGTGCTGAGCAGAGGTGGCCAAAGCAGATGCTGGGGAGTAC 660
Db 3720 AAGGCTGCACACGAAGGCTGTGTGCTGAGCAGAGGTGGCCAAAGCAGATGCTGGGGAGTAC 3779
Qy 661 AGCTGCGAGGCTGGGGGCCAGAGAGTCTCTTTTCAACTGSCATCAGAGAGCCCAAGGCA 720
Db 3780 AGCTGCGAGGCTGGGGGCCAGAGAGTCTCTTTTCAACTGSCATCAGAGAGCCCAAGGCA 3839
Qy 721 GTGTTTGCAGAGAGCAGTGGTGTGATTAATGAGGTGGGACATGAGGAGGGGCCAGTGCC 780
Db 3840 GTGTTTGCAGAGAGCAGTGGTGTGATTAATGAGGTGGGACATGAGGAGGGGCCAGTGCC 3899
Qy 781 ACATGAGCTGTGAGGTGGCCAGGCCAGACAGAGGTGACGTGTGTAAGGATGGGAAG 840
Db 3900 ACATGAGCTGTGAGGTGGCCAGGCCAGACAGAGGTGACGTGTGTAAGGATGGGAAG 3959
Qy 841 AAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGGCTGCGGCTGCATCGGCAGCTGGTG 900
Db 3960 AAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGGCTGCGGCTGCATCGGCAGCTGGTG 4019
Qy 901 GTGCAGCAGGAGGCGAGCAGTCTGGGAGTACACTGTGAGGCTGGGGGCCAGGCG 960
Db 4020 GTGCAGCAGGAGGCGAGCAGTCTGGGAGTACACTGTGAGGCTGGGGGCCAGGCG 4079
Qy 961 CTCTCTTCCACCTGGAGTGTTCAGAGCCCAAGGCGGTGT 1001
Db 4080 CTCTCTTCCACCTGGAGTGTTCAGAGCCCAAGGCGGTGT 4120

RESULT 4
US-10-988-426
; Sequence 426, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyun
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: NO. US20030219745A1el Nucleic Acids and
; FILE REFERENCE: Polypeptides
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/7774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt FL_genes Version 2.0
; SEQ ID NO 426
; LENGTH: 2768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)..(2738)
US-10-120-988-426

Query Match 24.8%; Score 248.4; DB 17; Length 2768;

Best Local Similarity 57.4%; Pred. No. 1.8e-58;
Matches 510; Conservative 0; Mismatches 366; Indels 12; Gaps 3;
Qy 75 GCACAGGCGGCAAGACGGATGCCGGGACTACAGCTGCGAGCGCGGAGGAGGT 134
Db 1874 GGAGGTGCCCGCCAGCTGACGCGGGCGCAGCTGAAAGTTCTGTGGCCAAACGCAATTGAGAG 1933
Qy 135 CTCCTTCGCGCTGCACATACACAGAGCCCAAGATGATGTTTCAAAAGGAGCAGTCAAGTCA 194
Db 1934 CAGCATCCGATGAGGTCCGGGCGGCCAGGGCTGACTGCCAACAGCCGCGCAGCCGC 1993
Qy 195 TAAAT-----GAGGTGACAGCTGAGCGGGGGCCAGTGCATGCTGAGCTGTGAGGTGGC 248
Db 1994 AGCTGCCCGGAGGTGCTGCTCGGCTGCACAGAGGCGCAGCTGTGGGTGAGCTGTC 2053
Qy 249 CCAGGCCACAGACGAGGTGACGTGTACAAGATGGGAAGAGCTGAGCTCCAGCTCAAA 308
Db 2054 AGATCAGGCTCGCGCTGTGAGCTGGCTGAAGATGGTGCACACTGTCCCAGGGCCCCAA 2113
Qy 309 AGTGGCATGAGGTCAAAAGGGTGCACACGAGAGCTGTGTGTCACAGGCGGGCAAAAGC 368
Db 2114 GTATGAGGTGACGACATCGGCCGGGCGGGTCTCTCTTGTGCGAGATGTGGCCCGGA 2173
Qy 369 AGATGCTGGGAGTACAGCTGTGAGGTGGGGGCCAGAGAGTCTCTTCCACTGCAAT 428
Db 2174 CGATGCAGGCTCTTACAGTGCCTCAGCCGCGGGGGCGCATCGCTACCCAGCTCTCGGT 2233
Qy 429 CACAGAGCCCAAGGGGTGTTTGCAGAGGACAGTCAAGTGCATAATGAGGTGCAAGCTGA 488
Db 2234 GCAAGGCTCTCGCGCTCTTTCGCAAGAGACATGGCGGCGAGCTGT---GTGGATGCCGT 2290
Qy 489 GCGCGGGACCACTGCCATGTGAGCTGTGAGGTGGCCAGCCCGCAGAGGAGTCAAGCTG 548
Db 2291 GGCTGGGGGCGCGCGAGTTTGAGTGTGAGACTCTCGGCTCTCTGAGCCCAAGCGTCACTG 2350
Qy 549 GTACAAGGACCGGAAGAGTGTGAGCTCCAGCTCAAAAGTACGATGAGGTCAAGGGCTG 608
Db 2351 GTACAAGGATGGCATGGAGCTGGGCCCATCCGGTGTGAGCGCTTCTTGCAGGAGGATGTGGG 2410
Qy 609 CACACG--AAGGCTGTGTGAGCAGGTGGGCAAGACAGATGCTGGGGAGTACAGCTG 665
Db 2411 GACCGGCAACCGCTGTGTGCGAGCCACAGTCAACAGGAGGATGAAGCACTACTCTG 2470
Qy 666 CGAGCTGGGGCCAGAGAGTCTCTTTCAACTGCATCAGAGCCCAAGGAGTGT 725
Db 2471 CCGGTGGGCGAGACTCTGTGGACTTCGGCTCCGGCTCTCTGAGCCCAAGGCGGTGT 2530
Qy 726 TGCCAAAGGAGGATGTGTGCAATATGAGTCCGACTGAGGCGAGGCGGCGAGTCCACACT 785
Db 2531 TGCCAAAGGAGCAGCCCGCTGTCAGGAGGTGTCAGGCTGAGGTGGGGGCGCGCCAGCT 2590
Qy 786 GAGCTGTGAGTGGCGCCAGGCCCAGACAGAGGTGACGTGGTGTACAGGATGGGAAGCT 845
Db 2591 GAGCTGTGAGTGGCGCCAGGACAGATGAGGTGACGTGTGTAAGGAGCGGAGAGT 2650
Qy 846 GAGCTCCAGTTCGAAAGTGGCATAGAGGTGCGGGCTGCATGCGGCGAGCTGGTGGTCA 905
Db 2651 GAGCTCCAGTTCGAAAGTGGCATGTCAGCTGGAGGCTGTGGGTGTATGCGGAGCTGTGGT 2710
Qy 906 GCAGGCGAGGCGAGCAGATGCTGGGGAGTACACTGTGAGGCTGGGG 953
Db 2711 GCAGGTGGCGCAGGACAGTCCCGGATAGTACAGCTGTGAAGCCAGGTG 2758

RESULT 5
US-10-027-632-272054
; Sequence 272054, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129


```
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 272054
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-272054

Query Match      18.5%; Score 185.6; DB 13; Length 578;
Best Local Similarity 87.5%; Pred. No. 3.1e-41;
Matches 203; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 153 CACAGAGCCCAAGATGATGTTTGCAGAGGAGCAGTCAGTGCATATATAGGTGCAGGCTGA 212
DB 347 CCCAGAGCCCAAGGTGGTGTTCAGAGGAGCAGCAGCAGAGGAGGTGCAGGCTGA 406

QY 213 GCGGGGGCCAGTGCCTGCTGAGCTGAGGTGGCCCGCCAGCCAGCGAGGTGACGTG 272
DB 407 GCGGGGGCCAGTGCCTGCTGAGCTGAGGTGGCCCGCCAGCCAGCGAGGTGACGTG 466

QY 273 GTACAAGGATGGAGAGCTGAGTCCAGCTCAAAAGTGGGCATGGAGGTCAAAAGGCTG 332
DB 467 GTACAAGGATGGAGAGCTGAGTCCAGCTCAAAAGTGGGCATGGAGGTGCGGTGGAGGCGGTGGGCTG 526

QY 333 CACACGAGGCTGTGCTGCACAGCGGGGCAAGCAGATGCTGGGGAGTAC 384
DB 527 CACACGAGGCTGTGCTGCACAGCGGGGCAAGCAGATGCTGGGGAGTAC 578

RESULT 6
US-10-027-632-272054
; Sequence 272054, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827,129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 272054
```

```
; LENGTH: 578
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-272054

Query Match      18.5%; Score 185.6; DB 17; Length 578;
Best Local Similarity 87.5%; Pred. No. 3.1e-41;
Matches 203; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 153 CACAGAGCCCAAGATGATGTTTGCAGAGGAGCAGTCAGTGCATATATAGGTGCAGGCTGA 212
DB 347 CCCAGAGCCCAAGGTGGTGTTCAGAGGAGCAGCAGCAGAGGAGGTGCAGGCTGA 406

QY 213 GCGGGGGCCAGTGCCTGCTGAGCTGAGGTGGCCCGCCAGCCAGCGAGGTGACGTG 272
DB 407 GCGGGGGCCAGTGCCTGCTGAGCTGAGGTGGCCCGCCAGCCAGCGAGGTGACGTG 466

QY 273 GTACAAGGATGGAGAGCTGAGTCCAGCTCAAAAGTGGGCATGGAGGTCAAAAGGCTG 332
DB 467 GTACAAGGATGGAGAGCTGAGTCCAGCTCAAAAGTGGGCATGGAGGTGCGGTGGAGGCGGTGGGCTG 526

QY 333 CACACGAGGCTGTGCTGCACAGCGGGGCAAGCAGATGCTGGGGAGTAC 384
DB 527 CACACGAGGCTGTGCTGCACAGCGGGGCAAGCAGATGCTGGGGAGTAC 578

RESULT 7
US-09-726-643-22
; Sequence 22, Application US/09726643
; Patent No. US20020028449A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/09/726,643
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-643-22

Query Match      12.3%; Score 123.2; DB 9; Length 1645;
Best Local Similarity 49.5%; Pred. No. 5.6e-24;
Matches 376; Conservative 0; Mismatches 378; Indels 6; Gaps 2;

QY 230 TGCTGAGCTGTGAGGTGGCCCGCCAGCGGAGGTGACGTGGTACAAAGATGGGAAGA 289
DB 591 TGCTGACTTGTGAGCTCTCAAGGTTGACTTCCCGGCNACCTGTGTACAAAGATGGGCGAGA 650

QY 290 AGCTGAGCTCCAGCTCAAAAGTGGGCATGGAGGTCAAAGGTGCACACGGAGGTGGTGC 349
DB 651 AGGTGGAGGAGAGCGAGTTGCTGGTGAAGATGGATGGGCGCAACACACCGTCTGTATCC 710

QY 350 TGCCACAGCGGGGCAAGAGCAGATGCTGGGGAGTACAGTGTGAGGCTGGGGCCAGAGAG 409
DB 711 TGCTGAGGCCAAAGTCCAGGACAGTGGCGGATTTGAGTGCAGGACAGAAAGGGGTCTCGG 770

QY 410 TCTCTTCCACCTGCACATACAGAGCCCAAGGGGTGTTTTCGAAAGAGCAGTCAAGTGC 469
DB 771 CCTTCTCGGGGTCACTGTCCAAAGTCTCTCCGTGCACATCG---TGAACCCCGAGAAC 827

QY 470 ATAATGAGGTGCAGGCTGAGCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGCCAGC 529
DB 828 ATGTGTTTCGTGCATGCCATTAATCTTCCGAGTGTGTCTGCTGCGCTGTGAGGTGGACCGAG 887

QY 530 CCCAGACGAGGTGACGTGGTGTACAAGGACGGGAAGAGCTGAGCTCCAGTCCAAAGTAC 589
```


Db 888 AGGACGCCCTGTGCGTTGGTACAAAGGACGGGCGAGGAGTGCATTCGTGG 947
Qy 590 GCATGGAGTCAAGGGCTGCACCAAGGCTGGTAGTCAGACAGGTGGCAAGCAGATG 649
Db 948 TGCTGGAGAAATGAGGGGCCCATCGCCCGCTGGTGTCTGCCGCCACCCCTCAGACG 1007
Qy 650 CTGGGAGTACAGCTGCGAGGCTGGGGCCAGAGAGTCTCCTTTCAAATGACATCACAG 709
Db 1008 GGGCGAGTTTCAGTGCCTGCTGAGATGAGTGTGCCTACTTCACTGTCAACATCACAG 1067
Qy 710 AGCCCAAGGCAAGTGTTCGCAAGAGCAGTGTGGTGCATAAATAGAGTGGGACTAGAGCAG 769
Db 1068 A---CGTCTCTCGTGGATCGGTATCCAGCGCAAGGTGTATGTGGCAGCGGTGCGCC 1124
Qy 770 GGGCCAGTGCACACTGAGCTGTGAGGTGGGCCAGGCCAGACAGAGTGCAGTGTGACA 829
Db 1125 TGGAGCGTGTGGTGTGACCTGTGAGCTATGCCGCCCTGGGCAAGAGTGGCTGGACCA 1184
Qy 830 AGGATGGGAAGAGCTGAGCTCCAGTTCGAAAGTGGCATAGAGGCTCGCGGCTGCATGC 889
Db 1185 AGGATGGAGAGGAGTGTGGAGAGCCCGCGCTGCTCTGCAGAAAGACACTGTCC 1244
Qy 890 GGCAGCTGGTGTGCAGCAGGAGGCCAGGCGAGATGCTGGGAGTACACCTGTGAGGCTG 949
Db 1245 GCGGCTGGTGTGCCCGCTGTCCAGCTCGAGGACTCCGGCGAGTACTTGTGTGAAATTG 1304
Qy 950 GGGGCCAGCGCTCTCCTTCCACCTGGATGTTTCAGAGCC 989
Db 1305 ACGATGAGTGGCGCTCTTCACTGTACCGTCACAGAGTGC 1344

RESULT 8

US-10-042-141-22
; Sequence 22, Application US/10042141
; Publication No. US20020183503A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/10/042,141
; CURRENT FILING DATE: 2002-01-11
; PRIOR APPLICATION NUMBER: 09/726,643
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-042-141-22

Query Match 12.3%; Score 123.2; DB 13; Length 1645;
Best Local Similarity 49.5%; Pred. No. 5.6e-24;
Matches 376; Conservative 0; Mismatches 378; Indels 6; Gaps 2;

Qy 230 TGCTGAGCTGTGAGTGGCCCGCCAGAGCGGAGGTGACGTGTGTTACAAAGATGGGAAGA 289
Db 591 TGCTGACTGTGAGCTCTCAAGGGTGGACTTCCCGGCAACCTGGTACAAAGATGGGCAGA 650
Qy 290 AGCTGAGCTCCAGCTCAAAAGTGGGCATGGAGGTCAAAAGGTGCACACGGAGGCTGTGTC 349
Db 651 AGGTGGAGGAGCGAGTGTGCTGGTGGTGAAGATGGATGGCGCGCAAAACACCGTCTGATCC 710
Qy 350 TGCCACACGGGGCAAGACAGATGCTGGGGAGTACAGCTGTGAGCTGTGGGGCCAGAGAG 409
Db 711 TGCTGTAGGCCCAAAAGTCCAGACAGTGTGGCGAGTTTGTGATGTCAGGACAGAGGGGTCTCGG 770
Qy 410 TCTCTTCCACCTGCACATCAAGAGCCCAAGGGGGTGTGTTGCGAAGAGGAGTCAAGTGC 469

Db 771 CCTTCTTCGGCGTCACTGTGCCAGATCCTCCCGTGCACATCG---TGGACCCCGAGAAC 827
Qy 470 ATAATAGGTGTCAGGCTGAGCGGGGACCACCTGCATCTGAGCTGTGAGGTGCCCCAGC 529
Db 828 ATGTGTTCTGTCATGCCATTAACCTTCGAGTGTGTCACTGCTGGCTGTGAGGTGACCGAG 887
Qy 530 CCACAGACGAGGTCACCTGGTACAAAGACGGGAAGAGCTCAGCTCCAGCTCAAAAAGTAC 589
Db 888 AGGAGCCCTGTGCGTTGGTACAAAGACGGGAGGAGGTGGAGGAGTGAATCTCGTGG 947
Qy 590 GCATGGAGGTCAAGGGCTGCACACAAAGGTGTGTGTCAGCAGGTGGGCAAAAGCAGATG 649
Db 948 TGCTGGAGAAATGAGGGGCCCATCGCGCGCTGTGTGCTGCCGCCACCCATCCCTCAGACG 1007
Qy 650 CTGGGAGTACAGCTGCGAGGCTGGGGCCAGAGAGTCTCCTTTCAACTGCACATCACAG 709
Db 1008 GGGGCGAGTTTCAGTGTGCTGCGTGGAGATGAGTGTGCTTACTTCACTGTCAACCATCACAG 1067
Qy 710 AGCCCAAGGCAAGTGTGTTGCCAAGGAGCAGATGTTGGTGCATAATAGGTGCGGACTCAGGCAG 769
Db 1068 A---CGTCTCTCTGCGATCGTGTATCCAGCGCAAGGTGTATGTGGCAGCCGTGCGCC 1124
Qy 770 GGGCAGTGCACACTGAGCTGTGAGTGTGGCCAGGCCACAGACAGAGGTGACGTGGTACA 829
Db 1125 TGGAGCGTGTGTGCTGACCTGTGAGCTATGCGCGGCTGTGGCGAGAGTGTGCTGGACCA 1184
Qy 830 AGGATGGGAAGAGCTCAGCTCCAGTTCGAAAGTGGCATAGAGGCTCGGGCTGCATGC 889
Db 1185 AGGATGGAGAGGAGTGTGGAGAGCCCGCGCTGTCTCTGCAGAAAGACACTGTTC 1244
Qy 890 GGCAGCTGGTGTGCAGCAGGCGCCAGGCGAGATGCTGGGGAGTACACCTGTGAGGCTG 949
Db 1245 GCGGCTGGTGTGCTGCCCGCTGCCAGCTCGAGGACTCCGGCGAGTACTTGTGTGAAATTG 1304
Qy 950 GGGGCCAGCGGCTCTCCTTCCACCTGGATGTTTCAGAGCC 989
Db 1305 ACGATGAGTGGCGCTCTCCTTCACTGTCAACCGTCACAGAGTGC 1344

RESULT 9

US-10-919-272-22
; Sequence 22, Application US/10919272
; Publication No. US20050010042A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/10/919,272
; CURRENT FILING DATE: 2004-08-17
; PRIOR APPLICATION NUMBER: US/09/726,643
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 22
; LENGTH: 1645
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-919-272-22

Query Match 12.3%; Score 123.2; DB 19; Length 1645;
Best Local Similarity 49.5%; Pred. No. 5.6e-24;
Matches 376; Conservative 0; Mismatches 378; Indels 6; Gaps 2;

Qy 230 TGCTGAGCTGTGAGTGGCCCGCCAGGCGGAGGTGAGCTGTGTACAAAGATGGGAAGA 289
Db 591 TGCTGACTGTGAGCTCTCAAGGGTGGACTTCCCGGCAACCTGTGTACAAAGATGGGCAGA 650
Qy 290 AGCTGAGCTCCAGCTCAAAAGTGGGCATGGAGGTCAAAAGGTGCACACGGAGGCTGTGTC 349

Db	651	AGGTGGAGGAGCGAGGTTGCTGTGTGTAAGATGGATGGCGCAAAACCCGCTGTGATCC	710
Qy	350	TGCCACAGGGGGCAAAAGCAGATCTCGGAGTACAGCTGTGAGGCTGGGGCCAGAGAG	409
Db	711	TGCTGTAGGCGCAAGTCCAGGACAGTGGCAGTTTGTGATGCAGGACAGAGGGGTCTCGG	770
Qy	410	TCTCTTCCACCTGCACATCACAGAGCCCCAAGGGGGTGTTTTGCAGAGGACAGTCACTGC	469
Db	771	CTTCTTTCGGCGTCACTGTCCAAGATCTCCCGTGCACATCG--TGGACCCCGGAGAAC	827
Qy	470	ATAATGAGGTGCAGGCTGAGCGGGGACCACTGCATGCTGTGAGCTGTGAGGTGCCCCAGC	529
Db	828	ATGTGTTGTGCATGCCATACTTCCBAGTGTGTCACTGCTGGCCTGTGAGGTGGACCGAG	887
Qy	530	CCACAGACGGAGGTGACGCTGGGTACAAAGGACGGGAAGAAAGCTGAGCTCCAGCTCAAAAGTAC	589
Db	888	AGGACGCCCTGTGCGTTGGTACAAAGACGGCGAGGAGGTGGAGGAGTGACTTCGTGG	947
Qy	590	GCAATGGAGGTCAAGGGTGCACACGAAGGCTGGTGTGTCACAGAGGTGGGCAAGCAGATG	649
Db	948	TGCTGGAGAATTAGGGGGCCCCATCGCGCCTGGTGTCTGCCCGCACCCATCCCTCAGACG	1007
Qy	650	CTGGGGAGTACAGCTGCGAGGCTGGGGCCACAGAGTCTCTTTTCACTGCACATCAACAG	709
Db	1008	GGGCGGAGTTTCAGTGCCTGCTGGAGATGAGTGTGCCTACTTCACTGTCAACATCAACAG	1067
Qy	710	AGCCCAAGGCAGTGTTTTCCAAAGGAGCAGTTGGTGCATAATAGAGTGGGACTGAGGCAG	769
Db	1068	A--CGTCTCTGTGTGATGTATCCAGCGCAAGGTGTATGTGGCAGCGTGGCC	1124
Qy	770	GGGCCATGGCCACTGAGCTGTGAGTGGCCAGGCCACACAGAGGTGACGTGGTACA	829
Db	1125	TGGAGCGTGTGGTGTCACTGTGAGCTATGTCGCGCCCTGGGACAGAGTGCCTGGACCA	1184
Qy	830	AGGATGGGAAGACGTGAGCTCCAGTTTCAAGTGGCCATAGAGCTTCGGGCTGCATGC	889
Db	1185	AGGATGGAGAGGAGGTGGTGGAGAGCCCCCGCTGTCTTGTGCAAGGAAGACACTGTCC	1244
Qy	890	GGCAGCTTGTGTGCAGCAGCGCCAGGCAGATGCTGGGGAGTACACCTGTGAGGCTG	949
Db	1245	GCCGCTGTGTGCCCCGTGTCCAGCTCGAGGACTCCGGCGAGTACTTGTGTGAATTG	1304
Qy	950	GGGGCCACGGGCTCTCTTCACCTGATGTTTCAGAGCC	989
Db	1305	ACGATAGTGGCCCTCTTCACTGTCACTGTCACCGTCAACAGATC	1344

RESULT 10

US-10-264-049-850
; Sequence 850, Application US/10264049
; Publication No. US20040005579A1

```

; GENERAL INFORMATION:
; APPLICANT: Birge et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 850
; LENGTH: 2170
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-264-049-850

```

Query Match 12.3%: Score 123.2: DB 17: Length 2170:

Best Local Similarity 49.5%; Pred. No. 5.8e-24; Score 123.2; DB 12.3%; Accuracy Match

Matches 376; Conservative 0; Mismatches 378; Indels 6; Gaps 2;

Qy	230	TGCTGAGCTGTGAGGTGGCCAGCGCCACAGCAGGAGTGCCTGGTCAACGGATGGGAAGA	289
Db	1082	TGCTGACTTGTGAGCTCTCAAGGGTGGACCTTCGCGGCAACTGGTCAACGAGATGGGCAGA	1141
Qy	290	AGCTGAGCTCCAGCTCAAAAGTGGCGATGAGAGTCAAAAGGGTCACACGAGGAGCTGTGC	349
Db	1142	AGGTGGAGGAGCGGAGTTGCTGGTGAAGATGAGTGGCGGCAACACCGTCTGATCC	1201
Qy	350	TGCCACAGCGGGCAAGACGAGATGCTGGGAGTACAGCTGTGAGGCTGGGGGCCAGAGAG	409
Db	1202	TGCTGAGGCCAAAGTCCAGGACAGTGGCGAGTTTGAGTGCAGACAGAAAGGGTCTCGG	1261
Qy	410	TCTCTCTCCACTGCGACATCAGAGAGCCAAAGGGGTGTTTGCAGAGGACGACTCAGTGC	469
Db	1262	CTTCTTTGGCGCTCACTGTCCAAGATCTCCCGTGACACATCG---TGGACCCCGGAGAAC	1318
Qy	470	ATAATGAGGTGCAGGCTGAGCGGGGACCACTCCCAATGCTGAGCTGTGAGGTGCGCCAGC	529
Db	1319	ATGTGTTGTGCAATGCCATNACTTCGAGTGTGTATGCTGGCTGTGAGGTGGACCGAG	1378
Qy	530	CCCAGACGGAGGTGACGTGGTGTACAAGGACGGGAAGAGCTGAGCTCCAGCTCAAAAGTAC	589
Db	1379	AGGACGCCCTGTGCTGTGTGTACAAGGACGGGACGAGGTGGAGGAGAGTGACTTCGTGG	1438
Qy	590	GCATGAGAGTCAAGGGCTGCACACGAAGGCTGTGTAGTCAGACAGGTGGGCAAGCAGATG	649
Db	1439	TGCTGGAGAATGAGGGGGCCCCATCGCGCGCTGTGTGCTGCCCGCACCCAGCCCTCAGACG	1498
Qy	650	CTGGGAGGTACAGCTCGAGGCTGGGGCCAGAGAGTCTCTTTTCAACTGSCATACACAG	709
Db	1499	GGGGCGAGTTTCAAGTGGCTGCCTGGAGATGAGTGTGCTACTTCACTGTCAACATACAG	1558
Qy	710	AGCCCCAAGGCAGTGTTTGCCAAGGAGCAGTTGCTGCTAATAGAGTGCAGCTCAGGCAG	769
Db	1559	A---GCTCTCTCTGTGATCGTATCCAGCGGCAAGGTGTATGTGGACCGCTGCGCC	1615
Qy	770	GGGCCAGTGCCACACTGAGCTGTGAGGTGGCCAGAGCCACAGACAGAGGTGACGTGTGTA	829
Db	1616	TGAGAGCGTGTGTGTGCTGACCTGTGAGCTATGCGCGGCCCTGGGCAGAGGTGCGCTGGACCA	1675
Qy	830	AGGATGGGAAGAGCTGAGCTCCAGTTTCGAAAGTGCAGATAGAGGCTGCGGGTGCATGC	889
Db	1676	AGGATGAGAGGAGGTGTGTGGAGAGCCCGCGCTGCTCTGTGCAAGGAAGACACTGTCTC	1735
Qy	890	GGCAGCTGTGTGTGCACAGCGCAGGCAGCAGATGCTGGGGAGTACACCTCTGACGGCTG	949
Db	1736	GCCGCTGTGTGCTGCCCGTGTCCAGCTCGAGGACTCCGGCGAGTACTTGTGTGAATTG	1795
Qy	950	GGGGCCAGCGGCTCTCTCTTCCACTCGATGTTTTCAGAGCC	999
Db	1796	ACGATGAGTCGGGCTCTCTTCACTGTCAACCTCACAGATC	1835

RESULT 11

US-10-093-463-73 ; Sequence 73, Application US/10093463 ; Publication No. US20030208039A1

```

: GENERAL INFORMATION:
:
: APPLICANT: Padigaru, Muralidhara
: APPLICANT: Shenoy, Suresh
: APPLICANT: Kekuda, Ramesh
: APPLICANT: Gusev, Vladimir
: APPLICANT: Pochart, Pascal
: APPLICANT: Zhong, Mei
: APPLICANT: Rastelli, Luca
: APPLICANT: Mezes, Peter
: APPLICANT: Smithson, Glennda
: APPLICANT: Guo, Xiaojia
: APPLICANT: Gerlach, Valerie
: APPLICANT: Casman, Scacie
: APPLICANT: Boldog, Ferenc
: APPLICANT: Li, Li

```


; NAME/KEY: misc feature
 ; OTHER INFORMATION: Incyte ID No: 7326129CB1
 US-10-476-397-15

Query Match	9.6%	Score 95.8	DB 18	Length 7564
Best Local Similarity	46.4%	Pred. No. 2.4e-16		
Matches 436	Conservative 0	Mismatches 482	Indels 21	Gaps 3
Qy	48	CAAAGGCTGCAGACGGAGGCTGGTGGTCAGCAGGACGGAAGACGGATGCCGGGACTA	107	
Db	1938	CNAGGGCCGGCCACACGCTGGTACTGAAGAGCATCCAGGGGCTGATCGGGCATAGT	1997	
Qy	108	CAGCTCGAGGCCNAGGGCCAGAGGGTCTCTTTCCGCTGCAATCAAGAGCCCAAGAT	167	
Db	1998	ACGCGCTCTCTCCCTGAAGGTGTGCACTCTGCCCGCTCGAGGTCCGAGTGAAGCCGT	2057	
Qy	168	GATGTTTGCRAAGAGAGCAGTCAGTGCATATAGGTGACGGCTCAGGCGGGGCCAGTGC	227	
Db	2058	GGTGTCTTGAAGCGCTG-----GATGACCTGTCCGAGAGAGCGGCGACCCCT	2108	
Qy	228	CATGCTGAGCTGTGAGTGGCCAGGCCACAGACGGAGGTACGCTGGTGAAGAATGGGAA	287	
Db	2109	GGCCCTCGAGTGTGAAGTCTCTGACCCGAGGCCATGTGTGTGGCGCAAGATGSCGT	2168	
Qy	288	GAGCTGAGCTCCAGCTCAAAAGTGGGCATGAGGTCAAAAGGTGACACGGAGGCTGGT	347	
Db	2169	GCAGCTGGGGCCCAAGTGAAGATATGATCTTCTTGCAACCGCGGGCACGCGGGGGCTCGT	2228	
Qy	348	GCTGCCACAGCGCGGCAAGACAGATGCTGGGGAGTACAGCTGTGAGGCTCGGGGCCAGAG	407	
Db	2229	GGTGCATGACGTGAGCCCTGAAGACGCGGCTGTACACCTGCCACATGGCTCCGAGGA	2288	
Qy	408	AGTCTCTTTCACCTGCACATCAGAGGCCAAGGGGGTGTTCGAGAGAGAGTCAAGT	467	
Db	2289	GACCCGGGCCGGGTCCGCGTGCACGATCTGCACGTGGGATCACCAGAGGGTG---AA	2345	
Qy	468	GCATAATGAGGTGCAGGCTCAGGCGGGACCACTGCCATCTGAGCTGTGAGTGGGCCCA	527	
Db	2346	GACAAATGAGGTGCTGGAAGGGAAAGCTGCAGCTTTGAGTGGCTCTGTCCCAAGAG	2405	
Qy	528	GCCCCAGACGGAGGTGACGTGGTCAAGGACGGGAAGAGCTGAGCTTCAGCTCAAAAGT	587	
Db	2406	TGCCAGCGACCCGGCCATGTGGACAGTTCGGTGGGAAGACAGTGGGAGCTCCAGCCGCTT	2465	
Qy	588	ACGCATGGAGGTCAAGGGCTGCACAGAAAGCTGGTAGTCAGCAGGTGGGCAAGCAGA	647	
Db	2466	CCAGGCCACAGTCAAGGCCCGAAATATCTTGGTGGTCCGGAGGCTGCACCAAGTGA	2525	
Qy	648	TGCTGGGAGTACAGCTGCGAGGCTGGGGCCACAGAGTCTCTTTCAACTGCACATCAC	707	
Db	2526	TGCCGGGGAGGTGGTCTTCTGTGCGGGGCTCACTCCAAAGGCTCACTCATTTGTGAG	2585	
Qy	708	AGAGCCCCAAGCAGTGTTCGCCAAGGACAGTGGTGGTCATAATGAGTGGGACTGAGGC	767	
Db	2586	AGAGAGCCCGGCCCATCATCAAGCCCTTGGAAAGCACAGTGGGTGGCGCCAGGGGAGGA	2645	
Qy	768	AGGGGCCAGTGCACACATGAGCTGTGAGGTGGGCCAGGCCCCAGACAGAGTGACTGGTA	827	
Db	2646	CGTGG-----GCTGCGCTGTGAGCTGTCACGGCGGGAAACGCCCTGTCACCTGGCT	2696	
Qy	828	CARAGGTGGAAAGACGTGACTCCAGTTCGAAAGTCCGCATAGAGGCTCGGGGCTGCAT	887	
Db	2697	GAAGGACAGAAAGGCCATCCGCAAGAGCCAGATATGATGTGGTCTGCGAGGGCACAT	2756	
Qy	888	CGGCGAGCTGGTGTGACGAGGCAGGCCAGGCAGATGCTGGGAGGTACACCTGTGAGGC	947	
Db	2757	GGCCATGCTGGTCATCCGCGGGGCTCGCTCAAGGACCGGGCGAGTACACGTGTGAGGT	2816	
Qy	948	TGGGGGCCAGGGGCTCTCTTCACCTCGGATGTTTCAGA	986	
Db	2817	GGAGGCTTCCAAGAGCACAGCCAGCTCCATGTGGAGA	2855	

```

; Patent No. US2002028449A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 26 Human secreted proteins
; FILE REFERENCE: P2040P1
; CURRENT APPLICATION NUMBER: US/09/726,643
; CURRENT FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/15187
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/137,725
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 190
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44
; LENGTH: 2254
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-726-643-44

Query Match      8.8%; Score 88.4; DB 9; Length 2254;
Best Local Similarity 50.4%; Pred.No. 2.3e-14;
Matches 269; Conservative 0; Mismatches 261; Indels 4; Gaps 2;

Qy      180  GGAGCAGTCAGTGCATAATCAGGTGCAGGCTGAGGCGGGGGCCAGTGCCTCATGCTGAGCTG 239
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1309  GGACCCCGAGACATGTGTTCTGTGCATGCCATACTTCGAGTGTGTCTATGCTGCGCTG 1368
Qy      240  TGAGGTGCGCCAGCGCCAGACGGAGGTGACGTGGTACAAGGATGGGAAGAAGCTGAGCTC 299
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1369  TGAGGTGGACCGAGAGGACGCCCTCTGCGTTGGTACAAGGACGGGACGAGGTGGAGGA 1428
Qy      300  CAGCTCAAAAGTGGGCATGGAGGTCAAGGGTGCACACGAGGCTGGTGTGCTGCCACAGGC 359
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1429  GATGTACTTCGTGGTGTGTGAGTAATGAGGGGCCCAATCGCCGCTGGTGTGCTGCCGCCAC 1488
Qy      360  GGGCAAAAGCAGATGCTGGGGAGTACAGCTGTGAGGCTGGGGGCGCAGAGAGTCTCTTCCA 419
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1489  CCAGCCTCAGACGGGGCGAGTTTCAGTGCCTGCTGTGAGATGAGTGTGCTTACTTCAC 1548
Qy      420  CTTGCACATCACAGAGCCCAAGGGGTGTTTGGCAAGGAGCAGTCAAGTGCATATGAGGT 479
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1549  TGTCAACATCACAGACGTCTCTCTGTGGATCGGTAT---CCAGCGGCAAGGTGTATGT 1605
Qy      480  GCAGGCTGAGCGGGGACCACTGCCATGCTGAGCTGTGAGGTGGGCCCGAGCCCGACACGGA 539
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1606  GGCAGCGTGGCGCTGGAGCGTGTGTGTCTGACCTGTGAGCTATGCGGCGCTTGGGCGAGA 1665
Qy      540  GGTGACGTGGTACAAGGACGGGAAGAGTGAAGTCCAGCTCCAGTCAAAAGTACGCAATGAGGT 599
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1666  GGTGCGCTGGACCAAGGATGGAGAGGAGGTG-GTGGAGACCCCGCGCTGCTCTCTGCAGAA 1724
Qy      600  CAAGGGCTGCACACGAGGCTGGTAGTGCAGCAGGTGGGCAAGCAGATGCTTGGGAGTA 659
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1725  GGAAGACACTGTCCGCGCGCTGGTGTGTCGCGCTGTCCAGCTCGAGGATCTCGGCGAGTA 1784
Qy      660  CAGCTCGAGGCTGGGGGCGCAGAGAGTCTCTTTCAACTGCACATCACAGAGCC 713
      |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      1785  CTTGTGTGAATTGACGATGAGTGCAGCTCGGCTCTCTTCACTGTCAACCTCACAGATC 1838

Search completed: March 21, 2005, 16:07:33
Job time : 633.052 secs

```

Search completed: March 21, 2005, 16:07:33
Job time : 633.052 secs

THIRTY FOUR EIGHT
BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:02:07 ; Search time 175.678 Seconds
(without alignments)
9323.380 Million cell updates/sec

Title: US-10-077-130-4_COPY_10500_11500
Perfect score: 1001
Sequence: 1 gcccctgtgagtgaggaa.....ttcatgaagatgtgaaaaa 1001

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	67.6	6.8	7218	1	US-08-232-463-14
2	46.4	4.6	767677	4	US-09-949-016-12147
3	46.4	4.6	767677	4	US-09-949-016-17361
C 4	42.4	4.2	1926	3	US-09-249-585A-4
C 5	42.4	4.2	1931	2	US-09-130-114-2
C 6	42.2	4.2	4276	4	US-09-949-016-4900
C 7	42.2	4.2	5065	4	US-09-949-016-744
C 8	42.2	4.2	50453	4	US-09-949-016-12486
C 9	42.2	4.2	51242	4	US-09-949-016-12486
C 10	41.6	4.2	767677	4	US-09-949-016-12147
C 11	41.6	4.2	767677	4	US-09-949-016-17361
12	41.2	4.1	289	3	US-09-007-005-17
13	41.2	4.1	289	3	US-09-244-796-17
14	40.8	4.1	601	4	US-09-949-016-61224
C 15	40.6	4.1	1166	3	US-09-072-596-323
C 16	40.6	4.1	1166	4	US-09-072-967-328
C 17	39.6	4.0	63930	4	US-09-949-016-12270
C 18	39.6	4.0	109690	4	US-09-949-016-13525
19	39.4	3.9	601	4	US-09-949-016-168436
20	39.4	3.9	43414	4	US-09-949-016-12839
21	39.4	3.9	43415	4	US-09-949-016-16491
C 22	39.2	3.9	965	1	US-08-736-361A-3
C 23	39.2	3.9	1576	1	US-08-736-361A-1
C 24	38.8	3.9	1975	4	US-09-308-345A-6
25	38.8	3.9	2295	4	US-09-949-016-5079
C 26	38.8	3.9	2456	3	US-08-813-150-5
C 27	38.8	3.9	2456	4	US-08-546-553-5

Sequence 93, Appl
Sequence 93, Appl
Sequence 177691,
Sequence 358, App
Sequence 349, App
Sequence 16821, A
Sequence 35, Appl
Sequence 31, Appl
Sequence 35, Appl
Sequence 6866, Ap
Sequence 580, App
Sequence 2, Appl
Sequence 3, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 14, Appl
Sequence 1, Appl

ALIGNMENTS

RESULT 1

US-08-232-463-14/C
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-Fls
US-08-232-463-14

Query Match

6.8%; Score 67.6; DB 1; Length 7218;

QY 340 AAATGTGAGCTCAGATTCTGTGCTGCTGCATG---GCAGACACTGGGGAGTACTCTGTC 396
Db |||||
QY 817 GAGGACGAGACCGGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGAC 758
Db |||||
QY 397 GTGTGCGGCGAGGAGGACCTCGGCTATGCTACCGTCTAGGGCTTACCATCAAGTTC 456
Db |||||
QY 757 GGGGAGGACGGGAGGAGGACGAGGACGGGAGGACGAGGACGGGAGGAGGACGAGGAC 698
Db |||||
QY 457 ACAGAGGGGTCTGAGGAACGAGAGCCACAGAGGGGCAACAGAGCGTCTCGGTGTGAG 516
Db |||||
QY 697 GACGAGGACGGGAGGACGGGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAC 638
Db |||||
QY 517 CTGACAGATGCCCCCTGTGAGTGTGGTGAAGGGCATGAGACCTCTCAGAGATGAGAC 576
Db |||||
QY 637 GGGGAGGACGGGAGGACGGGAGGACGAGGACGGGAGGAGGACGAGGACGGGAGGAC 578
Db |||||
QY 577 AGACACGCTGAGCAGGACGGGCGCAGGTGTGAGCTGAGATCGAGATCCGGGCTGTGGCA 636
Db |||||
QY 577 GGGGAGGACGGGAGGAGGACGAGGACGGGAGGACGGGAGGAGGAGGACGAGGACGGGAG 518
Db |||||
QY 637 GAGGACGCTGGGGAGTACTGTGCTATGTGCGGAAGGAGAGGACCTCAGCCATGCTCACC 696
Db |||||
QY 517 GAGGACGAGGACGGGAGGACGGGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 458
Db |||||
QY 697 GTCAGGGCCATGCTCCCAAGTTCTATAGAGGTCTGAGGAATGAAGAGGCCACAGAGGG 756
Db |||||
QY 457 GAGGAGGACGAGGACGGGAGGAGGACGAGGACGGGAGGAGGACGGGAGGAGGAGGAG 398
Db |||||
QY 757 GACAGGGCCACGCTGTGTGTGCTGAGCTGAGCAGGAGGACCGCTGAGTGTGAGGAGGG 816
Db |||||
QY 397 GACGAGGACGGGAGGAGGACGAGGACGGGAGGAGGACGGGAGGAGGAGGAGGAGGAG 338
Db |||||
QY 817 CATGACACCTCTCAGAGATGGGACAGACAGCCTGAGGACGAGGAGGAGGAGGAGGAG 864
Db |||||
QY 337 GAGGACGAGGACGGGAGGAGGACGAGGACGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 290
Db |||||

RESULT 6

US-09-949-016-4900/c
; Sequence 4900, Application US/09949016
; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 4900

LENGTH: 4276

TYPE: DNA

ORGANISM: Human

US-09-949-016-4900

Query Match 4.2%; Score 42.2; DB 4; Length 4276;

Best Local Similarity 50.8%; Pred. No. 0.65;

Matches 101; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 492 GGCAACAGCCGTCTGCGGTGTGAGCTGAGCAAGATGGCCCCCTGAGTGGTGGAGGG 551

Db |||||

QY 310 GCCGGCAGCGCGCTCCCTCCACTTTGATAGCGGTGTCCACCGAGGGCTGCTCGCAGGA 251

Db |||||

QY 552 GCATGAGACCTCTCAGAGATGAGACAGACAGCCTGAGGACGAGGACGAGGACGAGGAGTGA 611

Db |||||

QY 250 GGCGAGCGCGGTGCGGTGCTCTGCGGCGCGGCGGAGGCTGCGGGCTGCGCCGA 191

Db |||||

QY 612 GCTGCAGATCCGGGCCTCTGTGGCAGAGGACGCTGGGGAGTACCTGTGCATGTGCGGGA 671

Db 190 GGCGCCGGGGCTCTCGGGCTCGGCCCGCAGGCTTGGGGGGTCTCTGTCTCGCCGGGGGA 131

QY 672 GGAGAGGACCTCAGCCATG 690

Db 130 CGCGTCGCCCTCGGCGATG 112

RESULT 7

US-09-949-016-744/c

; Sequence 744, Application US/09949016

; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 744

LENGTH: 5065

TYPE: DNA

ORGANISM: Human

US-09-949-016-744

Query Match 4.2%; Score 42.2; DB 4; Length 5065;

Best Local Similarity 50.8%; Pred. No. 0.68;

Matches 101; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 492 GGCAACAGCCGTCTGCGGTGTGAGCTGAGCAAGATGGCCCCCTGAGTGGTGGAGGG 551

Db 1157 GCCGGCAGCGCGCTCCCTCCACTTTGATAGCGGTGTCCACCGAGGGCTGCTCGCAGGA 1098

QY 552 GCATGAGACCTCTCAGAGATGGAGACAGACAGCCTGAGGACGAGGACGGGCCAGGTGTA 611

Db 1097 GGCGGAGCGCGCTGCGGTGGCTGCTGGGCGCGCGCGGAGGCTGCGGGCTGCGCCGA 1038

QY 612 GCTGCAGATCCGGGCCTCTGTGCGCAGAGGACGCTGGGGAGTACCTGTGCATGTGCGGGA 671

Db 1037 GGCGCCGGGCGCTCGGGCTCGGCCCGCAGGCTTGGGGCGCTCTCTCTCGCCGGGGGA 978

QY 672 GGAGAGGACCTCAGCCATG 690

Db 977 CGCGTCGCCCTCGGCGATG 959

RESULT 8

US-09-949-016-16642/c

; Sequence 16642, Application US/09949016

; Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CL001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08


```
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16642
; LENGTH: 50453
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16642

Query Match          4.2%; Score 42.2; DB 4; Length 50453;
Best Local Similarity 50.8%; Pred. No. 1.4;
Matches 101; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 492 GGCACACGCGTGTGCGGTGAGCTGAGCAAGATGCCCCCGTGGAGTGGTGAAGGG 551
Db 2310 GCCGGCAGCCGCGCTCCCTCCACTTTGATAGCGGTGTCCACCGAGGGCTGCTGCAGGA 2251
Qy 552 GCATGAGACCTCAGAGATGAGAGACAGACAGCCTGAGGAGGACGGGGCCAGGTGTGA 611
Db 2250 GCGGAGGCGCGTGTGCTGTGGGGCGGCGCGGAGCTGCGGCTGCCCGGA 2191
Qy 612 GCTGAGATCCGGCGCTCTGTGGCAGAGGACGCTGGGGAGTACCTGTGCATGTGCGGAA 671
Db 2190 GCGCGCGGGGCGCTCGGGCTCGCGCGCCAGGCTTGGGGCGCTCTGTCTCTCGCGGGGA 2131
Qy 672 GGAGGAGACCTCAGCCATG 690
Db 2130 CGGTGCGCCCTCGGGGATG 2112

RESULT 9
US-09-949-016-12486/c
; Sequence 12486, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12486
; LENGTH: 51242
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12486

Query Match          4.2%; Score 42.2; DB 4; Length 51242;
Best Local Similarity 50.8%; Pred. No. 1.4;
Matches 101; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

Qy 492 GGCACACGCGTGTGCGGTGAGCTGAGCAAGATGCCCCCGTGGAGTGGTGAAGGG 551
Db 3157 GCCGGCAGCCGCGCTCCCTCCACTTTGATAGCGGTGTCCACCGAGGGCTGCTGCAGGA 3098
Qy 552 GCATGAGACCTCAGAGATGAGAGACAGACAGCCTGAGGAGGACGGGGCCAGGTGTGA 611
Db 3097 GCGGAGGCGCGTGTGCTGTGGGGCGGCGCGGAGCTGCGGCTGCCCGGA 3038
Qy 612 GCTGAGATCCGGCGCTCTGTGGCAGAGGACGCTGGGGAGTACCTGTGCATGTGCGGAA 671
Db 3037 GCGCGCGGGGCGCTCGGGCTCGCGCGCCAGGCTTGGGGCGCTCTGTCTCTCGCGGGGA 2978
Qy 672 GGAGGAGACCTCAGCCATG 690
Db 2977 CGGTGCGCCCTCGGGGATG 2959

RESULT 10
US-09-949-016-12147/c
; Sequence 12147, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12147
; LENGTH: 76767
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(76767)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12147

Query Match          4.2%; Score 41.6; DB 4; Length 76767;
Best Local Similarity 53.0%; Pred. No. 4.4;
Matches 89; Conservative 0; Mismatches 79; Indels 0; Gaps 0;

Qy 484 ACAGAAGGGGCAACAGCCGCTGCTGCGGTGTGAGCTGAGCAGATGGACACACAGCCCTGAGCAGGACGGGGCC 543
Db 494407 ACAGGAACTGGCGCTCGACATGGGGCAGGACGGGAGAGAGACACGGGGCAGACGGGGCAGGACGGGGAC 494348
Qy 544 TGAAGGGGCGATGAGACCCCTCAGAGATGGAGACAGACACAGCCCTGAGCAGGACGGGGCC 603
Db 494347 AGGAGGGGCGCAGGACAGGAGAGAGACACGGGGCAGACGGGGCAGGACGGGGAC 494288
Qy 604 AGGTGTAGCTGACAGTCCGGCGCTCTGTGCGCAGGACGCTGGGGAG 651
Db 494287 AGCGGGGCGCAGGACAGAGGGAGGACAGGGACGGGGCAGGAGACGGGG 494240

RESULT 11
US-09-949-016-17361/c
; Sequence 17361, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17361
; LENGTH: 76767
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
```


; LOCATION: (1)...(767677)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-949-016-17361

Query Match 4.2%; Score 41.6; DB 4; Length 767677;
Best Local Similarity 53.0%; Pred. No. 4.4;
Matches 89; Conservative 0; Mismatches 79; Indels 0;

484	ACAGAAGGGGCAACAGCGCGTCTCGGTGTGAGCTGAGCAAGATGCCCCCGCTGGAGTGG	543
QY		
494407	ACAGAGAATCTGGCCCTCGACATGGGCGCAGACGGGAGCAGGATGGGAACAGAGGGGGAC	494348
Db		
544	TGGAAGGGGCGATGAGACCCCTCAGAGATGGAGACAGACACAGCTGAGCGCACGCGGGCC	603
QY		
494347	AGGAGGGGCGAGGACAGGGGACAGAGAGAGACACGGGGCAGGACGGGGCAGAGCGGGAC	494288
Db		

Qy 604 AGGTGAGCTGCACATCCGGCGCTCGTGGCAGAGACGTTGGGGAG 651
||| ||| | | | | | | | | | | | |
Dδ 494287 AGCGGGGGCAGGACAGAGGGAGGACAGGGACGAGGAGAGACGGGGG 494240
||| ||| | | | | | | | | | | | |

RESULT 12
US-09-007-005-17
; Sequence 17, Application US/09007005B
; Patent No. 6258558

/ GENETIC INFORMATION:
 / APPLICANT: Szoatak, Jack W.
 / APPLICANT: Roberts, Richard W.
 / APPLICANT: Liu, Rihe
 / TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN

```

: TITLE OF INVENTION: FUSIONS
: FILE REFERENCE: 00786/350003
: CURRENT APPLICATION NUMBER: US/09/007,005B
: CURRENT FILING DATE: 1998-01-14
: EARLIER APPLICATION NUMBER: 60/035,963
: EARLIER FILING DATE: 1997-01-27
: EARLIER APPLICATION NUMBER: 60/064,491
: EARLIER FILING DATE: 1997-11-06
: NUMBER OF SEQ ID NOS: 33
: SOFTWARE: FASTSEQ for Windows Version 4.0
: SEQ ID NO 17
:

```

```

? LENGTH: 289
? TYPE: RNA
? ORGANISM: Artificial Sequence
? FEATURE:
? OTHER INFORMATION: Translation template
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)..(289)
? OTHER INFORMATION: n = A,T,C or G
? US-09-007-005-17

```

```
Query Match      4.1%; Score 41.2; DB 3; Length 289;
Best Local Similarity 5.2%; Pred.No. 0.51;
Matches 11; Conservative 98; Mismatches 103; Indels 0; Gaps 0;
```

[illegible][illegible]

RESULT 14
US-09-949

```

: Sequential
: Patent
: GENERAL
: APPLIC
: TITLE
: FILE R
: CURREN
: CURREN
: PRIOR
: PRIOR
: PRIOR
: PRIOR
: PRIOR
: PRIOR
: NUMBER
: SOFTWARE

```


; SEQ ID NO 61224
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-61224

Query Match
Best Local Similarity 4.1%; Score 40.8; DB 4; Length 601;
Matches 190; Conservative 1; Mismatches 208; Indels 5; Gaps 2;
QY 531 CCCCTGAGTGGTGAAGGGGCGATGAGACCCCTCAGAGATGAGACACAGACAGCCTGAG 590
DB 52 CCCCCCAGAGAGGAGGACTGGGACCAACAGACCCGCTGCGAGCAGCAAGGGCGTGGT 111
QY 591 CGAGACCGGGCCAGGTGTAGCTCAGATCCGCGGCTCGTGGCAGAGGACGCTGGGGA 650
DB 112 GTGCGCGGGGAGCAGCAGGAGACCCCTGAGCCCCACAGTCGCGGGCTCATGGCGCTGAGAA 171
QY 651 GTACCTGTGCATGTGCGGGAGGAGGACCTCAGCCATGCTCACCGTCAGGGCCATGCC 710
DB 172 GCATGCAAGAGGGCGGAGCCAGTAGGAGAGGAGCCGCAAGCGCAAGCATCCCGCAGACA 231
QY 711 TTCCAAGTTCATAGAGGGTCTGAGGAATGAAGAGGCCACAGAAAGGGGACACGGCCACGCT 770
DB 232 GTGTGAGTCT---AGCTGTGGGAGGCGGGGCTGGGTCAAGGGCCAGCGGCAGAGA 287
QY 771 GTGTGTGAGCTGAGCAAGCGGACACCGGTGAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAG 830
DB 288 GCATAGGGGCGCTGCG-GGGGCAGTGCCTAGGAGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 346
QY 831 AGATGGGACACACACACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 890
DB 347 CACTTTTGAATAGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 406
QY 891 CCTGCTGTGTGGATGCGGGGAGTACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 934
DB 407 TGAGAGGGTTCAGGAGGGCGGTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 450

RESULT 15

US-09-072-596-323/c
; Sequence 323, Application US/09072596
; Patent No. 6458366
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Campos-Neto, Antonia
; APPLICANT: Houghton, Raymond
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Twardzik, Daniel R.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Hendrickson, Ronald C.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF TUBERCULOSIS
; NUMBER OF SEQUENCES: 350
; CORRESPONDENCE ADDRESS:
; ADDRESS: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072.596
; FILING DATE: 05-MAY-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.

; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.417C9
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 323:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1166 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
US-09-072-596-323

Query Match
Best Local Similarity 4.1%; Score 40.6; DB 3; Length 1166;
Matches 157; Conservative 115; Mismatches 290; Indels 5; Gaps 3;
QY 276 GTGGAAGAAAGGGTCTGAGACCCCTCAGAGATGAGGACATACAGCCTGAGGACGAGCGG 335
DB 936 RKGRNGWAGMGAGATRAGWAGAGGATAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 877
QY 336 GACTAAATGTGAGCTGCAGATTCTGTGGCTCGCCATGGCAGACACTGGGGGAGTACTCTGTG 395
DB 876 GMTGKMTGGGGAGKNGTGTGTRRRYKNGNGAGGARMGAKGAKNGGAGGAGGAGGAGGAGGAG 817
QY 396 CGTGTGCGGGCAG-GAGAGGACCTCGGCTATGCTCACCGTCAAGGCTTACCCATCAAGT 454
DB 816 GWKRGMGKAGGAGAGKMGKMGAGTGAGARGGTCSMGNGANNNGMGNGTKAKRKKGAGA 757
QY 455 TCACAGAGGGTCTGAGGAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 511
DB 756 GAAGAANGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 697
QY 512 GTGAGCTGAGCAAGATGGCCCGCTGGAGTGGTGGAGGGGCGATGAGACCCCTCAGAGATG 571
DB 596 MTGTGGGGGSGWAGAGANMGTTMKGGMGWAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 637
QY 572 GAGACAGACACAGCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 631
DB 636 GMGRKGRWNGRGRAMMGTTKNTTGGMGAGAGRKNGMGRCGCCGGMGAGTTTGAKAGA 577
QY 632 TGGCAGAGGAGCCTGGGGAGTACCTGTGCATGTGCGGGAAGGAGGAGGAGGAGGAGGAGGAGGAG 691
DB 576 GSGKKTGNNARAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 517
QY 692 TCACCGTCAGGGCCATGCTTCCAAAGTTTCATAGAGGGTCTGAGGAATGAAGAGGAGGAGGAG 751
DB 516 -AGNGKTGAGRYNGSGGGGKKNAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 458
QY 752 AAGGGGACACGGCCACGCTGT 811
DB 457 GGTGKNTGAGGRCSCGKGGKGRATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 398
QY 812 AGGGGATGAGACCTCAGAGATGGG 838
DB 397 CNNGRAXAGRASACCKMKSAKTTSMGG 371

Search completed: March 21, 2005, 14:38:47
Job time : 182.678 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:02:07 ; Search time 175.678 Seconds
(without alignments)
9323.380 Million cell updates/sec

Title: US-10-077-130-4_COPY_22500_23500
Perfect score: 1001
Sequence: 1 ggcagtcagtgacactggc.....ctgtcgaagtcgtggacct 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	999.4	99.8	5207	4	US-09-858-664A-1
2	999.4	99.8	5207	4	US-10-274-978-1
3	999.4	99.8	5207	4	US-10-274-978-3
4	999.4	99.8	5207	4	US-10-697-263-1
5	999.4	99.8	5207	4	US-10-697-263-3
6	164.8	16.5	182	4	US-09-513-999C-29729
7	92	9.2	2157	4	US-09-949-016-1966
8	90.4	9.0	2105	4	US-09-949-016-195
9	90.4	9.0	2132	2	US-09-159-385-3
10	90.4	9.0	2132	3	US-09-186-277-3
11	84.2	8.4	1584	4	US-09-799-451-205
12	82.8	8.3	1282	2	US-08-878-989-12
13	82.8	8.3	1282	3	US-09-272-796-12
14	82.8	8.3	1282	4	US-09-016-434-953
15	82	8.2	1429	2	US-09-159-385-4
16	82	8.2	1429	3	US-09-186-277-4
17	78	7.8	1788	4	US-09-797-039-9
18	78	7.8	2046	4	US-09-797-039-7
19	78	7.8	15376	4	US-09-949-016-11937
20	78	7.8	15376	4	US-09-949-016-13708
21	73.2	7.3	8865	4	US-09-949-016-4238
22	73.2	7.3	8906	2	US-08-826-267-1
23	70.4	7.0	1458	4	US-09-230-896C-5
24	69.6	7.0	4935	2	US-08-631-097-3
25	69.6	7.0	5886	3	US-08-810-712-9
26	68.6	6.9	6034	4	US-09-949-016-711
27	68.6	6.9	6043	4	US-09-949-016-2582

28	68.6	6.9	6046	4	US-09-949-016-5830	Sequence 5830, Ap
29	67	6.7	2169	4	US-09-016-434-1147	Sequence 1147, Ap
30	67	6.7	7220	4	US-09-949-016-3586	Sequence 3586, Ap
31	63.4	6.3	678	4	US-09-841-683-4	Sequence 4, Appl
32	63.4	6.3	711	4	US-09-841-683-6	Sequence 6, Appl
33	63.4	6.3	1191	4	US-09-841-683-10	Sequence 10, Appl
34	63.4	6.3	1224	4	US-09-841-683-8	Sequence 8, Appl
35	63.4	6.3	1485	4	US-09-801-876B-1	Sequence 1, Appl
36	63.4	6.3	1485	4	US-10-254-869-1	Sequence 1, Appl
37	63.4	6.3	1485	4	US-10-667-442-1	Sequence 1, Appl
38	63.4	6.3	1675	4	US-09-841-683-12	Sequence 12, Appl
39	62.2	6.2	2673	3	US-09-428-711A-1	Sequence 1, Appl
40	62.2	6.2	3995	3	US-09-428-711A-13	Sequence 13, Appl
41	61.4	6.1	3228	4	US-09-579-664B-5	Sequence 5, Appl
42	61.4	6.1	3228	4	US-10-355-975A-5	Sequence 5, Appl
43	61.4	6.1	3503	4	US-09-823-038A-58	Sequence 58, Appl
44	61	6.1	2112	4	US-10-116-326-5	Sequence 5, Appl
45	61	6.1	2334	4	US-10-003-690-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1

US-09-858-664A-1
; Sequence 1, Application US/09858664A
; Patent No. 6482624
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP
; CURRENT APPLICATION NUMBER: US/09/858.664A
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-858-664A-1

Query Match	99.8%	Score 999.4;	DB 4;	Length 5207;
Best Local Similarity	99.9%	Pred. No. 9.6e-222;		
Matches 1000;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	GGCCAGTCAGTGACACCTGGCCCTGCCAGGTGTGAGGAGCAGCAGCGTGTCTTCATCTCTGCCACCTCAAG	60	
Db	3628	GGCCAGTCAGTGACACCTGGCCCTGCCAGGTGTGAGGAGCAGCAGCGTGTCTTCATCTCTGCCACCTCAAG	3687	
Qy	61	AGCAAGACGAGCGCCCTTGGAGGAGCAGCAGCGTGTCTTCATCTCTGCCACCTCAAG	120	
Db	3688	AGCAAGACGAGCGCCCTTGGAGGAGCAGCAGCGTGTCTTCATCTCTGCCACCTCAAG	3747	
Qy	121	AACCTTCAGCTTCTGACCATCTCTGGTGTGGTGTGAGGAGCAGCAGCGTGTCTTCATCTCTGCCACCTCAAG	180	
Db	3748	AACCTTCAGCTTCTGACCATCTCTGGTGTGGTGTGAGGAGCAGCAGCGTGTCTTCATCTCTGCCACCTCAAG	3807	
Qy	181	AGCGTGAGCAATCGCTGGGAGCAGTACCAACCGGGCGTCTCCGAGGAGCAGAGGC	240	
Db	3808	AGCGTGAGCAATCGCTGGGAGCAGTACCAACCGGGCGTCTCCGAGGAGCAGAGGC	3867	
Qy	241	CCCTCATCTTCGCGATATCGGGAGGTGTACCGGAGGTGTACCGGAGGTGTCTGCTGTC	300	
Db	3868	CCCTCATCTTCGCGATATCGGGAGGTGTACCGGAGGTGTACCGGAGGTGTCTGCTGTC	3927	
Qy	301	TGGAAGCCCGTGGAAATCTTACGGCCCTGTGACCTACATTTGTCAGTGCAGCTAGAGGC	360	
Db	3928	TGGAAGCCCGTGGAAATCTTACGGCCCTGTGACCTACATTTGTCAGTGCAGCTAGAGGC	3987	


```
QY 361 GGAGCTGGACCACTGGCTCCGACATCTTTTGACTGCTGTAACCTGACCAAGCAAGCTC 420
Db |||||
QY 3988 GGAGCTGGACCACTGGCTCCGACATCTTTTGACTGCTGTAACCTGACCAAGCAAGCTC 4047
Db |||||
QY 421 TCCCGGGTGGACCTTACACTTCCGACAGGATGTTGACGAGGCAAGATGGGTCCC 480
Db |||||
QY 4048 TCCCGGGTGGACCTTACACTTCCGACAGGATGTTGACGAGGCAAGATGGGTCCC 4107
Db |||||
QY 481 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGCGCCAGACCACTGGCTCTCAGGAG 540
Db |||||
QY 4108 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGCGCCAGACCACTGGCTCTCAGGAG 4167
Db |||||
QY 541 GAGAGCAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 600
Db |||||
QY 4168 GAGAGCAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 4227
Db |||||
QY 601 CAGATCCAGAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 660
Db |||||
QY 4228 CAGATCCAGAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 4287
Db |||||
QY 661 GCGCTGGCGCCCAAGATCATCCCTTACCAACCCCAAGGACAAGACAGTGTGCGCGAA 720
Db |||||
QY 4288 GCGCTGGCGCCCAAGATCATCCCTTACCAACCCCAAGGACAAGACAGTGTGCGCGAA 4347
Db |||||
QY 721 TACGAGGCGCTCAAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 780
Db |||||
QY 4348 TACGAGGCGCTCAAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 4407
Db |||||
QY 781 AGCCCCCGGACCTGCTGCTCATCTTGGAGCTGTGCTTGGGCGGAGCTGCTCCCTGC 840
Db |||||
QY 4408 AGCCCCCGGACCTGCTGCTCATCTTGGAGCTGTGCTTGGGCGGAGCTGCTCCCTGC 4467
Db |||||
QY 841 CTGGCGGAGAGGGCGCTTCTACTCAGAAATCCAGAGTGAAGGACTACTCTGCGCAGATGTTG 900
Db |||||
QY 4468 CTGGCGGAGAGGGCGCTTCTACTCAGAAATCCAGAGTGAAGGACTACTCTGCGCAGATGTTG 4527
Db |||||
QY 901 AGTGCCACCCAGTACTGTCGACCAACAGCAGATCTCTGCACTGGAGCTGAGTCCGAGAAC 960
Db |||||
QY 4528 AGTGCCACCCAGTACTGTCGACCAACAGCAGATCTCTGCACTGGAGCTGAGTCCGAGAAC 4587
Db |||||
QY 961 ATGATCATCACCAGATACAACTGCTCAAGTCTGAGGACCT 1001
Db |||||
QY 4588 ATGATCATCACCAGATACAACTGCTCAAGTCTGAGGACCT 4628
Db |||||
```

RESULT 2

```
US-10-274-978-1
; Sequence 1, Application US/10274978
; Patent No. 6670164
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV
; CURRENT APPLICATION NUMBER: US/10/274,978
; CURRENT FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Human
US-10-274-978-1
```

```
Query Match 99.8%; Score 999.4; DB 4; Length 5207;
Best Local Similarity 99.9%; Pred. No. 9.6e-22;
Matches 1000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 GGCGAGTCAAGTCACTGGCTCCGAGTGTACGCCAGCCAGCTGCCCCAGGCCACCTGG 60
Db |||||
QY 3628 GGCGAGTCAAGTCACTGGCTCCGAGTGTACGCCAGCCAGCTGCCCCAGGCCACCTGG 3687
Db |||||
QY 61 AGCAAGAGCGAGGCCCTCCCTGGAGAGCAGCAGCGTGTCTCTCATCTTCTGCCCACCTCAAG 120
Db |||||
QY 3688 AGCAAGAGCGAGGCCCTCCCTGGAGAGCAGCAGCGTGTCTCTCATCTTCTGCCCACCTCAAG 3747
Db |||||
QY 121 AACTTCCAGCTTCTGACATCTCTGGTGTGTGAGGACCTGGGTGTGTACACCTGC 180
Db |||||
QY 3748 AACTTCCAGCTTCTGACATCTCTGGTGTGTGAGGACCTGGGTGTGTACACCTGC 3807
Db |||||
QY 181 AGCGTGAAGCAATGCGCTGGGAGCAGTGAACACACGCGCGTCTCCGAGAGCAGAGCGC 240
Db |||||
QY 3808 AGCGTGAAGCAATGCGCTGGGAGCAGTGAACACACGCGCGTCTCCGAGAGCAGAGCGC 3867
Db |||||
QY 241 CCCTCATCTTTCGCAATCCCGGATATCGGGAAGGTGTACCGGATGGGTGCTCTGGTC 300
Db |||||
QY 3868 CCCTCATCTTTCGCAATCCCGGATATCGGGAAGGTGTACCGGATGGGTGCTCTGGTC 3927
Db |||||
QY 301 TGGAGGCCCGTGGAACTCTACGGCCCTGTGACCTTACATTTGTCAGTGCAGCTTAGAAGGC 360
Db |||||
QY 3928 TGGAGGCCCGTGGAACTCTACGGCCCTGTGACCTTACATTTGTCAGTGCAGCTTAGAAGGC 3987
Db |||||
QY 361 GGCGAGTCAAGTCACTGGCTCCGAGTGTACGCCAGCCAGCTGCCCCAGGCCACCTGG 420
Db |||||
QY 3988 GGCGAGTCAAGTCACTGGCTCCGAGTGTACGCCAGCCAGCTGCCCCAGGCCACCTGG 4047
Db |||||
QY 421 TCCCGGGTGGACCTTACACTTCCGACAGGATGTTGACGAGGCAAGATGGGTCCC 480
Db |||||
QY 4048 TCCCGGGTGGACCTTACACTTCCGACAGGATGTTGACGAGGCAAGATGGGTCCC 4107
Db |||||
QY 481 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGCGCCAGCACAAGTGTGAGGAG 540
Db |||||
QY 4108 TACAGCAGCCCTCGGAGCAAGTCTCTCGGAGGCGCCAGCACAAGTGTGAGGAG 4167
Db |||||
QY 541 GAGAGCAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 600
Db |||||
QY 4168 GAGAGCAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 4227
Db |||||
QY 601 CAGATCCAGAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 660
Db |||||
QY 4228 CAGATCCAGAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 4287
Db |||||
QY 661 GCGCTGGCGCCCAAGATCATCCCTTACCAACCCCAAGGACAAGACAGTGTGCGCGAA 720
Db |||||
QY 4288 GCGCTGGCGCCCAAGATCATCCCTTACCAACCCCAAGGACAAGACAGTGTGCGCGAA 4347
Db |||||
QY 721 TACGAGGCGCTCAAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 780
Db |||||
QY 4348 TACGAGGCGCTCAAGGGCGGTGACGCCCAACCCCTGCCAGCACAAGACCTTCGCAATTCAGACA 4407
Db |||||
QY 781 AGCCCCCGGACCTGCTGCTCATCTTGGAGCTGTGCTTGGGCGGAGCTGCTCCCTGC 840
Db |||||
QY 4408 AGCCCCCGGACCTGCTGCTCATCTTGGAGCTGTGCTTGGGCGGAGCTGCTCCCTGC 4467
Db |||||
QY 841 CTGGCGGAGAGGGCGCTTCTACTCAGAAATCCAGAGTGAAGGACTACTCTGCGCAGATGTTG 900
Db |||||
QY 4468 CTGGCGGAGAGGGCGCTTCTACTCAGAAATCCAGAGTGAAGGACTACTCTGCGCAGATGTTG 4527
Db |||||
QY 901 AGTGCCACCCAGTACTGTCGACCAACAGCAGATCTCTGCACTGGAGCTGAGTCCGAGAAC 960
Db |||||
QY 4528 AGTGCCACCCAGTACTGTCGACCAACAGCAGATCTCTGCACTGGAGCTGAGTCCGAGAAC 4587
Db |||||
QY 961 ATGATCATCACCAGATACAACTGCTCAAGTCTGAGGACCT 1001
Db |||||
QY 4588 ATGATCATCACCAGATACAACTGCTCAAGTCTGAGGACCT 4628
Db |||||
```

RESULT 3

```
US-10-274-978-3
; Sequence 3, Application US/10274978
; Patent No. 6670164
```


QY 361 GGCAGCTGGACCACTGGCCCTCCGACATCTTTTGACTGTCTGCTGACCAAGCAAGCTC 420
|
|
|
Db 3988 GGCAGCTGGACCACTGGCCCTCCGACATCTTTTGACTGTCTGCTGACCAAGCAAGCTC 4047
|
|
|
QY 421 TCCCGGGTGGACCTACACCTTCGCGAGGATGTGTGAGCAAGGCAAGGCAAGGCTCC 480
|
|
|
Db 4048 TCCCGGGTGGACCTACACCTTCGCGAGGATGTGTGAGCAAGGCAAGGCAAGGCTCC 4107
|
|
|
QY 481 TACAGCAGCCCTCGGAGCAAGTCTCTTGGAGGCGCCAGCCACTGGCTCTGAGGAG 540
|
|
|
Db 4108 TACAGCAGCCCTCGGAGCAAGTCTCTTGGAGGCGCCAGCCACTGGCTCTGAGGAG 4167
|
|
|
QY 541 GAGAGCCAGGCGGTGAGCCCAACCCCTGCCCCAGCAAAAGACCTTGGCAATTCAGACA 600
|
|
|
Db 4168 GAGAGCCAGGCGGTGAGCCCAACCCCTGCCCCAGCAAAAGACCTTGGCAATTCAGACA 4227
|
|
|
QY 601 CAGATCCAGAGGGCGCTTACAGCTGTGTGGGCAATGCTGGGAGAGGCGAGCGGGCG 660
|
|
|
Db 4228 CAGATCCAGAGGGCGCTTACAGCTGTGTGGGCAATGCTGGGAGAGGCGAGCGGGCG 4287
|
|
|
QY 661 GCGCTGGCGCGCAAGATCATCCCTACACCCCAAGGACAAGACAGCAGTGTGCGCGAA 720
|
|
|
Db 4288 GCGCTGGCGCGCAAGATCATCCCTACACCCCAAGGACAAGACAGCAGTGTGCGCGAA 4347
|
|
|
QY 721 TACGAGGCGCTCAAGGGCTGCGCCACCGCACCTGGCCAGCTGCACGCGACCTTACCTC 780
|
|
|
Db 4348 TACGAGGCGCTCAAGGGCTGCGCCACCGCACCTGGCCAGCTGCACGCGACCTTACCTC 4407
|
|
|
QY 781 AGCCCCCGCACCTGGTCTCATCTTGGAGCTGTCTTGGGCGGAGCTGTCTCCCTGC 840
|
|
|
Db 4408 AGCCCCCGCACCTGGTCTCATCTTGGAGCTGTCTTGGGCGGAGCTGTCTCCCTGC 4467
|
|
|
QY 841 CTGGCGGAGAGGGCTCTACTCAGAAATCCGAGGTGAAGGACTACTCTGGCAGATGTTG 900
|
|
|
Db 4468 CTGGCGGAGAGGGCTCTACTCAGAAATCTGAGGTGAAGGACTACTCTGGCAGATGTTG 4527
|
|
|
QY 901 AGTGCCACCCAGTACTGCACCAACAGCACATCTTGCACCTGGACCTGAGTCCGAGAAC 960
|
|
|
Db 4528 AGTGCCACCCAGTACTGCACCAACAGCACATCTTGCACCTGGACCTGAGTCCGAGAAC 4587
|
|
|
QY 961 ATGATCATCCGATACAACTGCTCAAGTCTGGACCT 1001
|
|
|
Db 4588 ATGATCATCCGATACAACTGCTCAAGTCTGGACCT 4628
|
|
|

RESULT 5

US-10-697-263-3
; Sequence 3, Application US/10697263
; Patent No. 6812014
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEROF
; FILE REFERENCES: CL000927-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,263
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 5207
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-697-263-3

Query Match

99.8%; Score 999.4; DB 4; Length 5207;

Best Local Similarity 99.9%; Pred. No. 9.6e-222;
Matches 1000; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GGCAGCTCAGTGACACTGGCTGCGAGGTGTGACCCAGCCAGCTGCCCCAGCCACCTGG 60
|
|
|
Db 3628 GGCAGCTCAGTGACACTGGCTGCGAGGTGTGACCCAGCCAGCTGCCCCAGCCACCTGG 3687
|
|
|
QY 61 AGCAAAAGCGAGAGCCCGCTTGGAGAGCAGCAGCGTGTCTCATCTCTGCCCACCTCAAG 120
|
|
|
Db 3688 AGCAAAAGCGAGAGCCCGCTTGGAGAGCAGCAGCGTGTCTCATCTCTGCCCACCTCAAG 3747
|
|
|
QY 121 AACTTCAGCTTCTGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
|
|
|
Db 3748 AACTTCAGCTTCTGACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3807
|
|
|
QY 181 AGCTGTAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
|
|
|
Db 3808 AGCTGTAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3867
|
|
|
QY 241 CCCTCATCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
|
|
|
Db 3868 CCCTCATCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3927
|
|
|
QY 301 TGAAGCCGCTGGAATCTTACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
|
|
|
Db 3928 TGAAGCCGCTGGAATCTTACGGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3987
|
|
|
QY 361 GGCAGCTGGAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
|
|
|
Db 3988 GGCAGCTGGAACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4047
|
|
|
QY 421 TCCCGGGTGGCACTACACTTCCGCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
|
|
|
Db 4048 TCCCGGGTGGCACTACACTTCCGCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4107
|
|
|
QY 481 TACAGCAGCCCTCGGAGCAAGTCTCTCTGCGAGGCGCCAGCCACTTGGCTCTGAGGAG 540
|
|
|
Db 4108 TACAGCAGCCCTCGGAGCAAGTCTCTCTGCGAGGCGCCAGCCACTTGGCTCTGAGGAG 4167
|
|
|
QY 541 GAGAGCCAGGCGGCTGAGCCCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
|
|
|
Db 4168 GAGAGCCAGGCGGCTGAGCCCAACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4227
|
|
|
QY 601 CAGATCCAGAGGGCGCTTTCAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
|
|
|
Db 4228 CAGATCCAGAGGGCGCTTTCAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4287
|
|
|
QY 661 GCGCTGGCGCGCAAGATCATCCCTACACCCCAAGGACAAGACAGCAGTGTGCTGCGGAA 720
|
|
|
Db 4288 GCGCTGGCGCGCAAGATCATCCCTACACCCCAAGGACAAGACAGCAGTGTGCTGCGGAA 4347
|
|
|
QY 721 TACAGGCGCTCAAGGGCTGCGCCACCGCACCTGCGCCAGCTGCACGCGACCTTACCTC 780
|
|
|
Db 4348 TACAGGCGCTCAAGGGCTGCGCCACCGCACCTGCGCCAGCTGCACGCGACCTTACCTC 4407
|
|
|
QY 781 AGCCCCCGCACCTGGTGTCTCATCTTGGAGCTGTCTTGGGCGGAGCTGTCTCCCTGC 840
|
|
|
Db 4408 AGCCCCCGCACCTGGTGTCTCATCTTGGAGCTGTCTTGGGCGGAGCTGTCTCCCTGC 4467
|
|
|
QY 841 CTGGCGGAGAGGGCTCTACTCAGAAATCCGAGGTGAAGGACTACTCTGGCAGATGTTG 900
|
|
|
Db 4468 CTGGCGGAGAGGGCTCTACTCAGAAATCTGAGGTGAAGGACTACTCTGGCAGATGTTG 900
|
|
|
QY 901 AGTGCCACCCAGTACTGCACCAACAGCACATCTTGCACCTGGACCTGAGTCCGAGAAC 960
|
|
|
Db 4528 AGTGCCACCCAGTACTGCACCAACAGCACATCTTGCACCTGGACCTGAGTCCGAGAAC 4587
|
|
|
QY 961 ATGATCATCCGATACAACTGCTCAAGTCTGGACCT 1001
|
|
|
Db 4588 ATGATCATCCGATACAACTGCTCAAGTCTGGACCT 4628
|
|
|

RESULT 6

QY 954 CGAGAACATGATCATCACCGAATACAACTGCTCAA 989
Db 519 GGAATAACATCATGCTGCTGGACAGACGTCGCCAA 554

RESULT 9

US-09-159-385-3
; Sequence 3, Application US/09159385
; Patent No. 5958748
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: PH-569
; CURRENT APPLICATION NUMBER: US/09/159,385
; CURRENT FILING DATE: 1998-09-23
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)..(1455)
US-09-159-385-3

Query Match 9.0%; Score 90.4; DB 2; Length 2132;
Best Local Similarity 58.0%; Pred. No. 1.8e-11;
Matches 160; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 714 GCGGAATACAGAGCCCTCAAGGGCCTGGCCACCGCAGCTGGCCCGCAGCGAGC 773
Db 279 GCGGAGGTGAACATCTCTGGGAGATCCGGACCCCAACATCATCACTTGCAGCAT 338
QY 774 CTACTCAGCCCCCGGACCTGGTGTCTCATCTTGGAGCTGTGCTCTGGCCCGGAGCTGCT 833
Db 339 CTTCGAGAACAAGACGGAGTGGTCTCATCTCTGGAGTGGTCTCTGGCGGGAGCTCTT 398
QY 834 CCCCTGCTGGCCGAGAGGGCTCTCTACTCAGATCCGAGTGAAGGACTACCTGTGGCA 893
Db 399 TGACTTCTCTGGGAGAAAGAGTCTGCTGACGGAGGACGAGGCCACCCAGTTCCTCAAGCA 458
QY 894 GATGTTGAGTGCACCCAGTACCTGCACACAGCACATCTCTGACCTGGACCTGAGGTC 953
Db 459 GATCTGACGGGCTTCACTACCTGCACTCTTAAGCGCATCGCACACTTTGACCTGAAGCC 518
QY 954 CGAGAACATGATCATCACCGAATACAACTGCTCAA 989
Db 519 GGAATAACATCATGCTGCTGGACAGACGTCGCCAA 554

RESULT 10

US-09-186-277-3
; Sequence 3, Application US/09186277
; Patent No. 6171841
; GENERAL INFORMATION:
; APPLICANT: AKIRA, SHIZUO
; APPLICANT: KAWAI, TARO
; TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
; FILE REFERENCE: 081356/0128
; CURRENT APPLICATION NUMBER: US/09/186,277
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: JP97/261589
; EARLIER FILING DATE: 1997-09-26
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 2132
; TYPE: DNA
; ORGANISM: Homo sapiens

; FEATURE:
; NAME/KEY: CDS
; LOCATION: (94)..(1455)
US-09-186-277-3

Query Match 9.0%; Score 90.4; DB 3; Length 2132;

Best Local Similarity 58.0%; Pred. No. 1.8e-11;
Matches 160; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
QY 714 GCGGAATACAGAGCCCTCAAGGGCCTGGCCACCGCAGCTGGCCCGCAGCGAGC 773
Db 279 GCGGAGGTGAACATCTCTGGGAGATCCGGACCCCAACATCATCACTTGCAGCAT 338
QY 774 CTACTCAGCCCCCGGACCTGGTGTCTCATCTTGGAGCTGTGCTCTGGCCCGGAGCTGCT 833
Db 339 CTTCGAGAACAAGACGGAGTGGTCTCATCTCTGGAGTGGTCTCTGGCGGGAGCTCTT 398
QY 834 CCCCTGCTGGCCGAGAGGGCTCTCTACTCAGAAATCCGAGGTGAAGGACTACCTGTGGCA 893
Db 399 TGACTTCTCTGGGAGAAAGAGTCTGCTGACGGAGGACGAGGCCACCCAGTTCCTCAAGCA 458
QY 894 GATGTTGAGTGCACCCAGTACCTGCACACAGCACATCTCTGACCTGGACCTGAGGTC 953
Db 459 GATCTGACGGGCTTCACTACCTGCACTCTTAAGCGCATCGCACACTTTGACCTGAAGCC 518
QY 954 CGAGAACATGATCATCACCGAATACAACTGCTCAA 989
Db 519 GGAATAACATCATGCTGCTGGACAGACGTCGCCNA 554

RESULT 11

US-09-799-451-205
; Sequence 205, Application US/09799451
; Patent No. 6783969
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Aeundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Xue, Aidong J.
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Ma, Yungqing
; APPLICANT: Yamazaki, Victoria
; APPLICANT: Chen, Rui-hong
; APPLICANT: Wang, Zhiwei
; APPLICANT: Wang, Dunrui
; APPLICANT: Yang, Yonghong
; APPLICANT: Wehrman, Tom
; APPLICANT: Ghosh, Reena
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6783969el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 803
; CURRENT APPLICATION NUMBER: US/09/799,451
; CURRENT FILING DATE: 2001-03-05
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: Pt_FL_genes Version 2.0
; SEQ ID NO 205
; LENGTH: 1584
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (51)..(1166)
US-09-799-451-205

Query Match 8.4%; Score 84.2; DB 4; Length 1584;
Best Local Similarity 54.3%; Pred. No. 4.5e-10;
Matches 195; Conservative 0; Mismatches 158; Indels 6; Gaps 1;

Qy	610	AGGGCGCGCTTCAGCGTGGTCGGCAATGCTGGGAGAAAGCCACGCGGGCGGGCGCTGGCC	6659
Db	204	AGGGGTGCTTCTCGAGGTGGTGTGGCTCAGGAGGGGGCTCCGACACCTCTGTGGCC	263
Qy	670	GCCAGATCATCCCCCTA-----CCACCCCAAGAGACAAGACAGCAGTGTCTGCGCGAATAC	723
Db	264	CTCAAGTGATCCCCCAAGAAGGCCCTCCGGGGCAAGGAGGCCCTGGTGGAGAACGAGATC	323
Qy	724	GAGGCCCTCAAGGGGCTTGGCCACCAGCCTGGGCCAGCTGCAAGCAGCGCTACCTCAGC	783
Db	324	GCAGTGCTCCGTAGGATCAGTCACCCCAACATCGTGGCTCTGGAGGATGTTCCACGAGGC	383
Qy	784	CCCCGACCTGGTGCTCATCTTGGAGCTGTGTCTCTGGGCCCGAGCTGTCTCCCTGTGCTG	843
Db	384	CTTCCACCTCTACTGGCCATGAACTGGTGAACGGGTGGCGAGCTGTTTTCACGCAATC	443
Qy	844	GCCGAGAGGGCTCTTACTCTCAGATCCGAGGTGAGGACTACTCTGTGGCGAGATGTTGAGT	903
Db	444	ATGAGCGGGCTCTCTACACAGAAAGATGCCAGCCATCTGTGGTGGGTTCAGGTCTTTGGC	503
Qy	904	GCCACCGATACCTGCACAAACAGACATCTCTGCACCTTGGACCTGGAGGTCCGAGAACAT	962
Db	504	GCGTCTCTTACTCTGACAGCTCGGGATCTGTGACCGGGACCTCAAGCCCGAAACCT	562

RECEIVED 12

US-08-878-989-12
; Sequence 12, Application US/08878989
; Patent No. 5885803

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lal, Preeti
APPLICANT: Goli, Surya K.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.

```

; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT06
; CLONE: 827431
US-08-878-989-12

```

Query Match 8.3%; Score 82.8; DB 2; Length 1282;
Best Local Similarity 52.5%; Pred. No. 8.9e-10;
Matches 208; Conservative 0; Mismatches 182; Indels 6

Qy	573	CAGCACAAGACTTTCGATTCACAGACACAGATCCAGAGGGCCGCTTCAGCGTGGTGGC	632
Db	230	CATCAGCAGCGGTCTACGAGATCCCGAGAGGCTCGGCTCGGGTGCTTCTCCGAGGTGGT	289
Qy	633	GCAATGCTGGGAAAGGCCACGCGGGCGGCTGGCCCAAGATCATCCCCCTA-----	686
Db	290	GCTGGCCAGAGAGCGGGGCTCCGCACACCTCGTGCCCTCAAGTGCATCCCCAAGAGGC	349
Qy	687	CCACCCCAAGACAAGACAGCAGTGCCTGCGCGAATAAGAGGCCCTCAAGGGCCTCGCCA	746
Db	350	CCTCGGGGCAAGGAGCCCTGGTGGAGACAGANTCGCAGTGCTCCGTAGGATCAGTCA	409
Qy	747	CCGCACTTGGCCAGCTGACGAGCGCTACTCAGCGCCCGGACCTGGTGCTCATCTTT	806
Db	410	CCCCAATCGTCGCTCTGGAGGATGTCCACGAGAGCCCTTCCACCTCTACCTGGCCAT	469
Qy	807	GGAGCTGTGCTCTGGGCCGAGCTGCTCCCTGCTGGCCAGAGGGCCTCCTACTCAGA	866
Db	470	GGAACTGGTGACGGGTGCGGAGCTGTTGACCGCATCATGAGCGCGGCTCCTACACAGA	529
Qy	867	ATCCGAGGTGAAGGACTACCTGTGGCAGATGTTGAGTGCACCCAGTACCTGCACAACCA	926
Db	530	GAAGGATCCAGCCATCTGGTGGGTGAGTGCTCTTGGCGCCGCTCTCTACCTGCACAGCT	589
Qy	927	GCATACCTTGACCTTGGACCTGAGGTCCGAGAACAT	962
Db	590	GGGGATCGTGTGCACCGGAGCTCTCAAGCCCGGAAACCT	625

RESULT 13

US-09-272-796-12
; Sequence 12, Application US/09272796
; Patent No. 6207148

GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Nell C.
APPLICANT: Guegler, Karl G.
APPLICANT: Lal, preeti
APPLICANT: Goli, Surya K.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSER: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA

NAME: Billings, Lucy J J
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 1282 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSNOT06
CLONE: 827431
US-09-272-796-12

Query Match 8.3%; Score 82.8; DB 3; Length 1282;
Best Local Similarity 52.5%; Pred. No. 8.9e-10;
Matches 208; Conservative 0; Mismatches 182; Indels 6; Gaps 1;

QY 573 CAGCACAAGACCTTCGCATTCACAGACACAGATCCAGAGGGCGCGCTTCAGCGTGTGCG 632
DB 230 CATCAGCAGCGTCTACGAGATCCGAGAGGCTCGGCTCGGCTGCTTCCTCGAGGTGGT 289

QY 633 GCAATGCTGGAGAGAGCGCGCGCGCGCTGCGCGCAAGATCATCCCTTA----- 686
DB 290 GCTGGCCAGGAGCGGGCTCGCACACCTCGTGGCCCTCAAGTCATCCCAAGAAGGC 349

QY 687 CCACCCCAAGGACAGACAGAGTGTGCGGGAATACAGAGGCCCTCAAGGGCTCGCGCA 746
DB 350 CCTCGGGGCAAGGAGGCGCTGTCGAGAAACGAGATCGCAGTGTCTCGTAGGATCAGTCA 409

QY 747 CCCGCACTTGGCCAGCTGCAAGCAGCTTACCTCAGGCCCGCGCACCTGTGCTCATCTT 806
DB 410 CCCCAACATCTGCTGCTGAGGATGTCACAGAGAGCCCTTCCCACTCTACCTGGCCAT 469

QY 807 GGAGCTGTGCTTGGGCGCGAGCTCTCCCTGCTGCGCGAGAGGGCTCTCTACTCAGA 866
DB 470 GGAACCTGTGAGCGGTGCGAGCTGTTGACCGCATCATGAGCGCGCTCTCTACACAGA 529

QY 867 ATCCAGGTGAGGACTACCTGTGCGAGATGTTGAGTGCACCCAGTACCTGCACACCA 926
DB 530 GAAGATGCGAGCCATCTGTGGTGGTCAAGTCTTGGCGCGCTCTCTACTGCAAGCCT 589

QY 927 GCACATCTGCACTTGGACCTGAGGTCCGAGAACAT 962
DB 590 GGGGATCTGCACCGGAGCTCAAGCCCGAAACCT 625

RESULT 14
US-09-016-434-953
Sequence 953, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 953:
SEQUENCE CHARACTERISTICS:
LENGTH: 1282 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: PROSNOT06
CLONE: 827431
US-09-016-434-953

Query Match 8.3%; Score 82.8; DB 4; Length 1282;
Best Local Similarity 52.5%; Pred. No. 8.9e-10;
Matches 208; Conservative 0; Mismatches 182; Indels 6; Gaps 1;

QY 573 CAGCACAAGACCTTCGCATTCACAGACACAGATCCAGAGGGCGCGCTTCAGCGTGTGCG 632
DB 230 CATCAGCAGCGTCTACGAGATCCGAGAGGCTCGGCTCGGCTGCTTCCTCGAGGTGGT 289

QY 633 GCAATGCTGGAGAGAGCGCGCGCGCGCTGCGCGCAAGATCATCCCTTA----- 686
DB 290 GCTGGCCAGGAGCGGGCTCCCGCACACCTCGTGGCCCTCAAGTCATCCCAAGAAGGC 349

QY 687 CCACCCCAAGGACAGACAGAGTGTGCTGCGCGAATACAGAGGCCCTCAAGGGCTCGCGCA 746
DB 350 CCTCGGGGCAAGGAGGCGCTGTCGAGAACGAGATCGCAGTGTCTCGTAGGATCAGTCA 409

QY 747 CCCGCACTTGGCCAGCTGCAAGCAGCTTACCTCAGGCCCGCGCACCTGTGCTCATCTT 806
DB 410 CCCCAACATCTGCTGCTGAGGATGTCACAGAGAGCCCTTCCCACTCTACCTGGCCAT 469

QY 807 GGAGCTGTGCTTGGGCGCGAGCTCTCCCTGCTGCGCGAGAGGGCTCTCTACTCAGA 866
DB 470 GGAACCTGTGAGCGGTGCGAGCTGTTGACCGCATCATGAGCGCGCTCTCTACACAGA 529

QY 867 ATCCAGGTGAGGACTACCTGTGCGAGATGTTGAGTGCACCCAGTACCTGCACACCA 926
DB 530 GAAGATGCGAGCCATCTGTGGTGGTCAAGTCTTGGCGCGCTCTCTACTGCAAGCCT 589

QY 927 GCACATCTGCACTTGGACCTGAGGTCCGAGAACAT 962
DB 590 GGGGATCTGCACCGGAGCTCAAGCCCGAAACCT 625

RESULT 15
US-09-159-385-4
Sequence 4, Application US/09159385
Patent No. 5958748
GENERAL INFORMATION:
APPLICANT: AKIRA, SHIZUO
APPLICANT: KAWAI, TARO
TITLE OF INVENTION: DNA CODING FOR SERINE/THREONINE KINASE
FILE REFERENCE: PH-569
CURRENT APPLICATION NUMBER: US/09/159,385
CURRENT FILING DATE: 1998-09-23
EARLIER APPLICATION NUMBER: JP97/261589
EARLIER FILING DATE: 1997-09-26
NUMBER OF SEQ ID NOS: 8


```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1429
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (10)..(1353)
US-09-159-385-4
```

```
Query Match      8.2%; Score 82; DB 2; Length 1429;
Best Local Similarity 58.7%; Pred. No. 1.4e-09;
Matches 142; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 740 TCGGCCACCCGACCTGGCCCCAGCTGACAGCAGCCCTACCTAGCCCCCGGCACCTGGTGC 799
Db      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 800 TCATCTTGGAGTGTGCTCTGGGCCGAGCTGCTCCCTGCTGGCCGAGAGGGCCTCCT 859
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 860 ACTCAGAAATCCGAGGTGAAGGACTACCTGTGGCAGATGTTGAGTGCCACCCAGTACCTGC 919
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 920 ACAACCCAGACATCTGTGACCTGGACCTGAGGTCCGAGAAATGATCATCACCGAATACA 979
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 980 AC 981
Db      |||
Qy 461 AC 462
Db      |||
```

Search completed: March 21, 2005, 14:38:50
Job time : 178.678 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:02:07 ; Search time 175.678 Seconds
(without alignments)
9323.380 Million cell updates/sec

Title: US-10-077-130-6_COPY_2000_3000
Perfect score: 1001
Sequence: 1 cgcctaccatccggagggtg.....ctgaggcaggggccaatgcc 1001

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405588

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	830.8	83.0	2768	4	US-09-774-528-426
C 2	64	6.4	1926	3	US-09-249-585A-4
C 3	64	6.4	1931	2	US-09-130-114-2
4	62	6.2	1651	4	US-09-902-540-3255
5	62	6.2	1954	4	US-09-902-540-1150
6	52.8	5.3	2695	4	US-09-620-312D-473
7	51.6	5.2	3575	4	US-09-023-655-1395
C 8	51.4	5.1	690	4	US-09-902-540-7374
C 9	51.4	5.1	2855	4	US-09-902-540-1935
10	51.4	5.1	18916	4	US-09-902-540-1057
C 11	51	5.1	7218	1	US-08-232-463-14
12	49.4	4.9	3766	4	US-09-949-016-1651
13	49.4	4.9	3778	4	US-09-949-016-359
14	49.4	4.9	50937	3	US-09-428-517-1
15	49	4.9	2670	4	US-09-902-540-8525
C 16	49	4.9	9321	4	US-09-902-540-898
17	48.4	4.8	4285	3	US-09-040-774-1
18	48.2	4.8	1539	3	US-08-582-740-69
19	48.2	4.8	1539	3	US-09-109-879-69
20	48.2	4.8	1598	4	US-09-023-655-915
21	48.2	4.8	1624	3	US-08-582-740-67
22	48.2	4.8	1624	3	US-09-109-879-67
23	48	4.8	5207	4	US-09-858-664A-1
24	48	4.8	5207	4	US-10-274-978-1
25	48	4.8	5207	4	US-10-274-978-3
26	48	4.8	5207	4	US-10-697-263-1
27	48	4.8	5207	4	US-10-697-263-3

```

28 47.8 4.8 2668 2 US-08-461-775-11 Sequence 11, Appl
29 47.8 4.8 2668 3 US-09-031-606-11 Sequence 11, Appl
30 47.4 4.7 3397 4 US-09-949-016-2188 Sequence 2188, Ap
31 47 4.7 4403765 3 US-09-103-840A-2 Sequence 2, Appli
32 47 4.7 4411529 3 US-09-103-840A-1 Sequence 1, Appli
33 46.8 4.7 2199 4 US-09-902-540-6388 Sequence 6388, Ap
34 46.8 4.7 2585 4 US-09-902-540-424 Sequence 424, App
35 46.4 4.6 1320 4 US-09-252-991A-165 Sequence 165, App
36 46.4 4.6 1686 4 US-09-252-991A-155 Sequence 155, App
37 46.4 4.6 1803 4 US-09-252-991A-170 Sequence 170, App
38 46.4 4.6 2610 4 US-09-252-991A-175 Sequence 175, App
39 46.2 4.6 1620 2 US-08-461-775-10 Sequence 10, Appl
40 46.2 4.6 1620 3 US-09-031-606-10 Sequence 10, Appl
41 45.8 4.6 603 4 US-09-902-540-9070 Sequence 9070, Ap
42 45.8 4.6 7719 4 US-09-902-540-969 Sequence 969, App
43 45.8 4.6 77536 4 US-09-410-551B-1 Sequence 1, Appli
44 45.8 4.6 77536 4 US-09-940-316B-1 Sequence 1, Appli
45 45.6 4.6 3114 4 US-09-902-540-2501 Sequence 2501, Ap

```

ALIGNMENTS

```

RESULT 1
US-09-774-528-426
; Sequence 426, Application US/09774528
; Patent No. 6743619
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Zhou, Ping
; APPLICANT: Goodrich, Ryle
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Zhang, Jie
; APPLICANT: Zhao, Qing A.
; APPLICANT: Yang, Yonghong
; APPLICANT: Xue, Aidong J.
; APPLICANT: Wehrman, Tom
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6743619el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802
; CURRENT APPLICATION NUMBER: US/09/774,528
; CURRENT FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_Fl_genes Version 2.0
; SEQ ID NO 426
; LENGTH: 2768
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (72)..(2738)
US-09-774-528-426

```

```

Query Match 83.0%; Score 830.8; DB 4; Length 2768;
Best Local Similarity 95.3%; Pred. No. 4.9e-173;
Matches 856; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
QY 1 CGCTACCATCCGGAGGTGCCCCCAGCTGCACGGGGCGCAGCTGAAGTTCGTGGCCA 60
Db 1861 CGCTACCATCCGGAGGTGCCCCCAGCTGCACGGGGCGCAGCTGAAGTTCGTGGCCA 1920
QY 61 ACGGCATTCAGACAGCATCCGGATGAGGTCCCGGGCGGGCGGCTGACTGCCACA 120
Db 1921 ACGGCATTCAGACAGCATCCGGATGAGGTCCCGGGCGGGCGGCTGACTGCCACA 1980
QY 121 AGCCGCCAGCGCAGCTGCCCGGGAGGTGCTGGCTCGCTCGCTCAGAGGAGCGCAGTGC 180
Db 1981 AGCCGCCAGCGCAGCTGCCCGGGAGGTGCTGGCTCGCTCAGAGGAGCGCAGTGC 2040

```


181 TGGCTGAGCTGTGATCAGATCAGGCTGCGCTGTGACGTGGCTGAAGGATGGTCCACACTGT 240
182 |||||
183 |||||
184 |||||
185 |||||
186 |||||
187 |||||
188 |||||
189 |||||
190 |||||
191 |||||
192 |||||
193 |||||
194 |||||
195 |||||
196 |||||
197 |||||
198 |||||
199 |||||
200 |||||
201 |||||
202 |||||
203 |||||
204 |||||
205 |||||
206 |||||
207 |||||
208 |||||
209 |||||
210 |||||
211 |||||
212 |||||
213 |||||
214 |||||
215 |||||
216 |||||
217 |||||
218 |||||
219 |||||
220 |||||
221 |||||
222 |||||
223 |||||
224 |||||
225 |||||
226 |||||
227 |||||
228 |||||
229 |||||
230 |||||
231 |||||
232 |||||
233 |||||
234 |||||
235 |||||
236 |||||
237 |||||
238 |||||
239 |||||
240 |||||
241 |||||
242 |||||
243 |||||
244 |||||
245 |||||
246 |||||
247 |||||
248 |||||
249 |||||
250 |||||
251 |||||
252 |||||
253 |||||
254 |||||
255 |||||
256 |||||
257 |||||
258 |||||
259 |||||
260 |||||
261 |||||
262 |||||
263 |||||
264 |||||
265 |||||
266 |||||
267 |||||
268 |||||
269 |||||
270 |||||
271 |||||
272 |||||
273 |||||
274 |||||
275 |||||
276 |||||
277 |||||
278 |||||
279 |||||
280 |||||
281 |||||
282 |||||
283 |||||
284 |||||
285 |||||
286 |||||
287 |||||
288 |||||
289 |||||
290 |||||
291 |||||
292 |||||
293 |||||
294 |||||
295 |||||
296 |||||
297 |||||
298 |||||
299 |||||
300 |||||
301 |||||
302 |||||
303 |||||
304 |||||
305 |||||
306 |||||
307 |||||
308 |||||
309 |||||
310 |||||
311 |||||
312 |||||
313 |||||
314 |||||
315 |||||
316 |||||
317 |||||
318 |||||
319 |||||
320 |||||
321 |||||
322 |||||
323 |||||
324 |||||
325 |||||
326 |||||
327 |||||
328 |||||
329 |||||
330 |||||
331 |||||
332 |||||
333 |||||
334 |||||
335 |||||
336 |||||
337 |||||
338 |||||
339 |||||
340 |||||
341 |||||
342 |||||
343 |||||
344 |||||
345 |||||
346 |||||
347 |||||
348 |||||
349 |||||
350 |||||
351 |||||
352 |||||
353 |||||
354 |||||
355 |||||
356 |||||
357 |||||
358 |||||
359 |||||
360 |||||
361 |||||
362 |||||
363 |||||
364 |||||
365 |||||
366 |||||
367 |||||
368 |||||
369 |||||
370 |||||
371 |||||
372 |||||
373 |||||
374 |||||
375 |||||
376 |||||
377 |||||
378 |||||
379 |||||
380 |||||
381 |||||
382 |||||
383 |||||
384 |||||
385 |||||
386 |||||
387 |||||
388 |||||
389 |||||
390 |||||
391 |||||
392 |||||
393 |||||
394 |||||
395 |||||
396 |||||
397 |||||
398 |||||
399 |||||
400 |||||
401 |||||
402 |||||
403 |||||
404 |||||
405 |||||
406 |||||
407 |||||
408 |||||
409 |||||
410 |||||
411 |||||
412 |||||
413 |||||
414 |||||
415 |||||
416 |||||
417 |||||
418 |||||
419 |||||
420 |||||
421 |||||
422 |||||
423 |||||
424 |||||
425 |||||
426 |||||
427 |||||
428 |||||
429 |||||
430 |||||
431 |||||
432 |||||
433 |||||
434 |||||
435 |||||
436 |||||
437 |||||
438 |||||
439 |||||
440 |||||
441 |||||
442 |||||
443 |||||
444 |||||
445 |||||
446 |||||
447 |||||
448 |||||
449 |||||
450 |||||
451 |||||
452 |||||
453 |||||
454 |||||
455 |||||
456 |||||
457 |||||
458 |||||
459 |||||
460 |||||
461 |||||
462 |||||
463 |||||
464 |||||
465 |||||
466 |||||
467 |||||
468 |||||
469 |||||
470 |||||
471 |||||
472 |||||
473 |||||
474 |||||
475 |||||
476 |||||
477 |||||
478 |||||
479 |||||
480 |||||
481 |||||
482 |||||
483 |||||
484 |||||
485 |||||
486 |||||
487 |||||
488 |||||
489 |||||
490 |||||
491 |||||
492 |||||
493 |||||
494 |||||
495 |||||
496 |||||
497 |||||
498 |||||
499 |||||
500 |||||
501 |||||
502 |||||
503 |||||
504 |||||
505 |||||
506 |||||
507 |||||
508 |||||
509 |||||
510 |||||
511 |||||
512 |||||
513 |||||
514 |||||
515 |||||
516 |||||
517 |||||
518 |||||
519 |||||
520 |||||
521 |||||
522 |||||
523 |||||
524 |||||
525 |||||
526 |||||
527 |||||
528 |||||
529 |||||
530 |||||
531 |||||
532 |||||
533 |||||
534 |||||
535 |||||
536 |||||
537 |||||
538 |||||
539 |||||
540 |||||
541 |||||
542 |||||
543 |||||
544 |||||
545 |||||
546 |||||
547 |||||
548 |||||
549 |||||
550 |||||
551 |||||
552 |||||
553 |||||
554 |||||
555 |||||
556 |||||
557 |||||
558 |||||
559 |||||
560 |||||
561 |||||
562 |||||
563 |||||
564 |||||
565 |||||
566 |||||
567 |||||
568 |||||
569 |||||
570 |||||
571 |||||
572 |||||
573 |||||
574 |||||
575 |||||
576 |||||
577 |||||
578 |||||
579 |||||
580 |||||
581 |||||
582 |||||
583 |||||
584 |||||
585 |||||
586 |||||
587 |||||
588 |||||
589 |||||
590 |||||
591 |||||
592 |||||
593 |||||
594 |||||
595 |||||
596 |||||
597 |||||
598 |||||
599 |||||
600 |||||
601 |||||
602 |||||
603 |||||
604 |||||
605 |||||
606 |||||
607 |||||
608 |||||
609 |||||
610 |||||
611 |||||
612 |||||
613 |||||
614 |||||
615 |||||
616 |||||
617 |||||
618 |||||
619 |||||
620 |||||
621 |||||
622 |||||
623 |||||
624 |||||
625 |||||
626 |||||
627 |||||
628 |||||
629 |||||
630 |||||
631 |||||
632 |||||
633 |||||
634 |||||
635 |||||
636 |||||
637 |||||
638 |||||
639 |||||
640 |||||
641 |||||
642 |||||
643 |||||
644 |||||
645 |||||
646 |||||
647 |||||
648 |||||
649 |||||
650 |||||
651 |||||
652 |||||
653 |||||
654 |||||
655 |||||
656 |||||
657 |||||
658 |||||
659 |||||
660 |||||
661 |||||
662 |||||
663 |||||
664 |||||
665 |||||
666 |||||
667 |||||
668 |||||
669 |||||
670 |||||
671 |||||
672 |||||
673 |||||
674 |||||
675 |||||
676 |||||
677 |||||
678 |||||
679 |||||
680 |||||
681 |||||
682 |||||
683 |||||
684 |||||
685 |||||
686 |||||
687 |||||
688 |||||
689 |||||
690 |||||
691 |||||
692 |||||
693 |||||
694 |||||
695 |||||
696 |||||
697 |||||
698 |||||
699 |||||
700 |||||
701 |||||
702 |||||
703 |||||
704 |||||
705 |||||
706 |||||
707 |||||
708 |||||
709 |||||
710 |||||
711 |||||
712 |||||
713 |||||
714 |||||
715 |||||
716 |||||
717 |||||
718 |||||
719 |||||
720 |||||
721 |||||
722 |||||
723 |||||
724 |||||
725 |||||
726 |||||
727 |||||
728 |||||
729 |||||
730 |||||
731 |||||
732 |||||
733 |||||
734 |||||
735 |||||
736 |||||
737 |||||
738 |||||
739 |||||
740 |||||
741 |||||
742 |||||
743 |||||
744 |||||
745 |||||
746 |||||
747 |||||
748 |||||
749 |||||
750 |||||
751 |||||
752 |||||
753 |||||
754 |||||
755 |||||
756 |||||
757 |||||
758 |||||
759 |||||
760 |||||
761 |||||
762 |||||
763 |||||
764 |||||
765 |||||
766 |||||
767 |||||
768 |||||
769 |||||
770 |||||
771 |||||
772 |||||
773 |||||
774 |||||
775 |||||
776 |||||
777 |||||
778 |||||
779 |||||
780 |||||
781 |||||
782 |||||
783 |||||
784 |||||
785 |||||
786 |||||
787 |||||
788 |||||
789 |||||
790 |||||
791 |||||
792 |||||
793 |||||
794 |||||
795 |||||
796 |||||
797 |||||
798 |||||
799 |||||
800 |||||
801 |||||
802 |||||
803 |||||
804 |||||
805 |||||
806 |||||
807 |||||
808 |||||
809 |||||
810 |||||
811 |||||
812 |||||
813 |||||
814 |||||
815 |||||
816 |||||
817 |||||
818 |||||
819 |||||
820 |||||
821 |||||
822 |||||
823 |||||
824 |||||
825 |||||
826 |||||
827 |||||
828 |||||
829 |||||
830 |||||
831 |||||
832 |||||
833 |||||
834 |||||
835 |||||
836 |||||
837 |||||
838 |||||
839 |||||
840 |||||
841 |||||
842 |||||
843 |||||
844 |||||
845 |||||
846 |||||
847 |||||
848 |||||
849 |||||
850 |||||
851 |||||
852 |||||
853 |||||
854 |||||
855 |||||
856 |||||
857 |||||
858 |||||
859 |||||
860 |||||
861 |||||
862 |||||
863 |||||
864 |||||
865 |||||
866 |||||
867 |||||
868 |||||
869 |||||
870 |||||
871 |||||
872 |||||
873 |||||
874 |||||
875 |||||
876 |||||
877 |||||
878 |||||
879 |||||
880 |||||
881 |||||
882 |||||
883 |||||
884 |||||
885 |||||
886 |||||
887 |||||
888 |||||
889 |||||
890 |||||
891 |||||
892 |||||
893 |||||
894 |||||
895 |||||
896 |||||
897 |||||
898 |||||
899 |||||
900 |||||
901 |||||
902 |||||
903 |||||
904 |||||
905 |||||
906 |||||
907 |||||
908 |||||
909 |||||
910 |||||
911 |||||
912 |||||
913 |||||
914 |||||
915 |||||
916 |||||
917 |||||
918 |||||
919 |||||
920 |||||
921 |||||
922 |||||
923 |||||
924 |||||
925 |||||
926 |||||
927 |||||
928 |||||
929 |||||
930 |||||
931 |||||
932 |||||
933 |||||
934 |||||
935 |||||
936 |||||
937 |||||
938 |||||
939 |||||
940 |||||
941 |||||
942 |||||
943 |||||
944 |||||
945 |||||
946 |||||
947 |||||
948 |||||
949 |||||
950 |||||
951 |||||
952 |||||
953 |||||
954 |||||
955 |||||
956 |||||
957 |||||
958 |||||
959 |||||
960 |||||
961 |||||
962 |||||
963 |||||
964 |||||
965 |||||
966 |||||
967 |||||
968 |||||
969 |||||
970 |||||
971 |||||
972 |||||
973 |||||
974 |||||
975 |||||
976 |||||
977 |||||
978 |||||
979 |||||
980 |||||
981 |||||
982 |||||
983 |||||
984 |||||
985 |||||
986 |||||
987 |||||
988 |||||
989 |||||
990 |||||
991 |||||
992 |||||
993 |||||
994 |||||
995 |||||
996 |||||
997 |||||
998 |||||
999 |||||
1000 |||||

RESULT 2

US-09-249-585A-4/c
; Sequence 4, Application US/09249585A
; Patent No. 6417002
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert
; TITLE OF INVENTION: METHOD FOR MAINTENANCE AND SELECTION OF EPISOMES
; FILE REFERENCE: 0867/0905
; CURRENT APPLICATION NUMBER: US/09/249,585A
; CURRENT FILING DATE: 1999-02-11
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1926
; TYPE: DNA
; ORGANISM: Epstein Barr Virus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(1926)
; OTHER INFORMATION: template strand of EBNA-1 DNA
US-09-249-585A-4

Query Match

6.4%; Score 64; DB 3; Length 1926;

Best Local Similarity 43.9%; Pred. No. 7.1e-05;
Matches 387; Conservative 0; Mismatches 480; Indels 15; Gaps 2;

QY 14 GGAGGTGCCGCCAGCTGACCGGGGCGAGCTGAAAGTTTCGTGTCACCAACGCAATGAGAG 73
DB 1127 GCAGGTGTGAGAGGGGGACCGAGAAAGTCTTGAAGGGGGGACCGAGAAAGTGCAGGA 1068
QY 74 CAGCATCCGGATGAGATCCGGGGCGGCCCGAGGAGCTGGGGCCGGAGGTGATGAGGAGCTGGGGCC 1008
DB 1067 GATGGGGCCCGCGAGGTGATGAGGAGCTGGGGCCGGAGGTGATGAGGAGCTGGGGCC 1008
QY 134 AGCTGCCCGGAGGTGCTGGCTCGGCTGCACGAGGAGCGCAGCTGTGGCTGAGCTGTCTC 193
DB 1007 GGAGTGCACGAGGAGCTGGGGCCGGAGGTGAGGAGCGAGGACGGGGAGGACGAGGACGG 948
QY 194 AGATCAGGCTCGGCTGTGAGCTGGCTGTAAGAGATGTTGACACTGTGTCACCAACGCAATGAGAG 253
DB 947 GGAGGAGCGAGGACGGGGGAGGACCGGGGAGGACGAGGACCGGGGAGGACGAGGAGGACG 888
QY 254 GTATGAGGTGTCAGGATCGGCTGAGTGCCTGAGCTGCGCTTACCGCTTACCGCTTCTCCGT 373
DB 887 GGAACGGGAGGACCGGGGAGGACGAGGACCGGGGAGGACGAGGACCGGGGAGGACGAGGAC 768
QY 374 GCAAGGCTTCGCGGCTTCTGACAGGACATGCGGGGAGGACGAGGACCGGGGAGGACGAGGAC 433
DB 767 GGAACGGGAGGACCGGGGAGGACGAGGACCGGGGAGGACGAGGACCGGGGAGGACGAGGAC 720
QY 434 TGGGGGCGCGGCTGAGTGTGAGACTCCGAGGACCGCTCCGAGGACCGCTCCGAGCTGACCTGTA 493
DB 719 CGGGGAGGACGAGGACCGGGGAGGACGAGGACCGGGGAGGACCGGGGAGGACCGGGGAGGAC 660
QY 494 CAAGGATGGCATGG---AGCTGGGCGCACTCCGCTGAGCGCTTCTTTCAGGAGGATGTGGG 550
DB 659 CGAGGACCGGGGAGGACGAGGACCGGGGAGGACCGGGGAGGACCGGGGAGGACGAGGACGAGG 600
QY 551 GACGCGGACCGGCTGGTGGCAGCACAGTCCAGGACGAGGACCGGGGAGGACGAGGACCGGGGAGG 610
DB 599 GGAAGGACGAGGACCGGGGAGGACCGGGGAGGACCGGGGAGGACGAGGACCGGGGAGGACGAG 540
QY 611 CCAGCTGGGCGAGGACTCTGTGAGCTTCCGGCTTCCGGCTTCTCTGAGCCCAAGGTGGTGT 670
DB 539 CGAGGAGGACGAGGACCGGGGAGGACGAGGACCGGGGAGGACCGGGGAGGACCGGGGAGGAC 480
QY 671 TGCTAAGGACGAGCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 730
DB 479 GGACGAGGACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 731 GAGCTGCGAGGTGGCGCCAGGCGCCAGGAGGTGAGTGTGTAACAAGGATGGGAAGAGCT 790
DB 419 GGAACGGGAGGACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 360
QY 791 GAGCTCCAGCTCGAAAGTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 850
DB 359 GGAAGGACCGGGGAGGACCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
QY 851 GCAAGGACCGGCGGAGGATGCCGGGAGGTATAGCTCCGAGGC 892
DB 299 CGAGGACCGGGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 258

RESULT 3

US-09-130-114-2/c
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Damaj, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes


```

; TITLE OF INVENTION: From Multiple Transfected Episoemes
;
; FILE REFERENCE: 0867/10903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1931
; TYPE: DNA
; ORGANISM: EBNA
US-09-130-114-2

```

Query Match	6.4%	Score 64	DB 2	Length 1931
Best Local Similarity	43.9%	Pred. No. 7.1e-05		
Matches 387	Conservative 0	Mismatches 480	Indels 15	Gaps 2
Qy	14	GGAGGTCCCGCCAGCCTGCACCGGGCGCAGCTGAAGTTCTGTGCCCAACGGCATTGAGAG	73	
Db	1127	GCAGGTCTGTGAGAGGGGACCGAGAAAGTGCTGAAGGGGGACCCAGAAAGTGTGACAGGA	1068	
Qy	74	CAGCATCCGGATGAGAGTTCGGGCGGCCCGAGGGCTGACTGCCAAAGCCCGCAGCCGC	133	
Db	1067	GATGGGCGCCCGAGGTGATGGAGAGAGCTGGGGCCGAGGTGATGGAGAGCTGGGGCC	1008	
Qy	134	AGTGTCCCGGAGGTGTGTGGCTCGGTGTGCACAGAGAGCGCAGCTGTGGCTGAGCTGTCT	193	
Db	1007	GGAGGTGACGAGGAGCTGGGGCCGAGGTGTGAGGACAGGACCGGAGGACGAGGACGG	948	
Qy	194	AGATCAGGCTCGGGCTGTGACGTGGCTGAAGGATGGTGCACACTGTCCCGAGGCCCAA	253	
Db	947	GGAGGAGGACGAGGACCGGGAGGACCGGGAGGACGAGGACCGGGAGGACCGGAGGACGA	888	
Qy	254	GTATGAGGTGCAGGCATCGGCCGGGCGGGGTCTCTTGTGCGAGATGTGGCCCGGGGA	313	
Db	887	GGACCGGGAGGACCGGGAGGACGAGGACCGGGAGGACCGGGAGGACGAGGACCGGGGA	828	
Qy	314	CGATGCAAGGCTCTTACAGTGTGCTACGCCGGGGGCGCATCGCTTACAGCTCTCCGT	373	
Db	827	GGACCGGGAGGAGGACGAGGACCGGGAGGACCGGGAGGACGAGGACCGGGAGGACGGGA	768	
Qy	374	GCAAGGCTCGCGGCTTCTGTACAAAGGACATGGCGGCAGCTGTGTGATGCCGTGGC	433	
Db	767	GGACGAGGACGGGG-----AGGACGGGGAGGACGAGGACCGGGAGGACGAGGA	720	
Qy	434	TGGGGGCGCGCGCAGTTTGTGTGAGACCTCCGAAGCCACCTCCACGTGCATCTGTA	493	
Db	719	CGGGGAGGACGAGGACCGGGAGGACGAGGACCGGGAGGACCGGGAGGACCGGGAGGAGGA	660	
Qy	494	CAAGGATGGCATGG---AGCTGGGCCACTCGGTGAGCGCTTCTTTCAGGAGGATGTGGG	550	
Db	659	CGAGGACGGGGAGGACGAGGACCGGGAGGACCGGGAGGACCGGGAGGACGAGGACCGGGGA	600	
Qy	551	GACCGGACCGGCTGTGTGGCAGCCACAGTCAACAGGAGGATGAAGGCACCTTACTCTTG	610	
Db	599	GGAGGACGAGGACCGGGAGGACCGGGAGGACCGGGAGGAGGACGAGGACCGGGAGGACGG	540	
Qy	611	CCGCGTGGGCGAGGACTGTGTGACTTCTCGGCTCCGCTCTCTCAGCCCAAGGTGTGTGT	670	
Db	539	GGAGGAGGACGAGGACCGGGAGGAGGACGAGGACCGGGAGGACCGGGAGGACCGGGAGGA	480	
Qy	671	TGCTAAGGAGCAGCTGGCACGCAGGAAGCTGTGAGGAGGAGGAGGCCACAGTGCACACT	730	
Db	479	GGACGAGGACCGGGAGGACCGGGAGGAGGACGAGGACCGGGAGGAGGACGAGGACCGGGGA	420	
Qy	731	GAGCTGCAGGTGGCCACGGCCAGACGAGAGGTGACGTGGTACAAGGATGGGAAGAAGCT	790	
Db	419	GGAACGGGGAGGACCGGGAGGAGGACGAGGACCGGGAGGAGGACGAGGACCGGGAGGACGG	360	
Qy	791	GAGCTCCAGCTCGAAAGTGTGCTGGAAGGCCACAGGCTGCACGCGCAGGCTGGTTGTGCA	850	
Db	359	GGAGGACCGGGAGGACCGGGAGGAGGAGGACGAGGACCGGGAGGAGGACGAGGACCGGGAGGA	300	
Qy	851	GCAGGCAGGCCAGCCGAGTGCAGGGGAGTATAGCTGCGAGGC	892	

Db
299 CGAGGACGGGGAGGGCGAGGACGAGGACGAGGACAAGGTGGC 258

RESULT 4

```

US-09-902-540-3255
; Sequence 3255, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 3255
; LENGTH: 1651
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-902-540-3255

```

Query Match 6.2%; Score 62; DB 4; Length 1651;
Best Local Similarity 48.1%; Pred. No. 0.00019;
Matches 176; Conservative 0; Mismatches 190; Indels

Qy	318	GCAGGCCTCTACAGTGGCGTCAGCCGGGGGGCCGATCGCCTACACAGCTCTCCGTGC	3177
Db	1211	GCGGCCGTCGAAGAGGGCGTCTGCTTGGCGGGCGGCTTACATCCGCTGCCTCAAG	1270
Qy	378	GGCCTCGCGCGCTTTCTGCACAAGGACATGGCGGGCAGCTGTGTGATGCGCGTGGCTGGG	437
Db	1271	GCCTGGAAGCGCCTTGACAGCTGTCCGGCGGTGAAGATTGGCGTGGACATCATCCGCCGC	1330
Qy	438	GGCCCGCGCGAGTTTGAGTGTGAGACCTCCGAAGCCCAAGTCCACGTCGACATCGGTACAAG	497
Db	1331	GCCGTGAGAGAGCCCTCGCCAGATCGTCGGCAACGGCGGCTCGAGGGCAGCGTGGTG	1390
Qy	498	GATGGCATTGAGAGCTGGGCCATCTCGGTGAGCGTCTTTCGACGAGAGATGTGGGAGCGCG	557
Db	1391	GTGAACAAGGTTCAAGAGAGTCTCCCGTCCGTTTCGGGTTCAACGCCCAAGGACCTAC	1450
Qy	558	CACCGGCTGTGGCAGCCACAGTCACAGGACAGATGAAGGCACCTACTCTCCCGCGTG	617
Db	1451	GAGGACCTGCTGCGCCCGCGGTATCGACCCGGCCAGGTGACGGCGCACCGCGCTGCAG	1510
Qy	618	GGCGAGGACTCTGTGGACTTTCGGGCTCCGCGTCTCTGAGGCCCAAGGTGGTGTGTTTGC	677
Db	1511	AACGCGGCTCCGTGTCTCGCTGATGCTCACCACCGAGGCCATGTTGGCGGAGCGTCCG	1570
Qy	678	GAGCAG 683	
Db	1571	AAGGAG 1576	

DEPT. T. 6

```

RESULI 5
US-09-902-540-1150
; Sequence 1150, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883

```


Query Match	Best Local Similarity	Score	DB 4	Length	19954
Matches 176; Conservative	0	Mismatches 190; Indels	0	Gaps	0
318	GCAGGCTCTACAGTGCCTGCTGACGCGCGGGGCGGATCGCTACCGTCTCCGTCGCA	377			
8811	CGCGCGCTGCAAGAGGCGTCTGCTCGCGGGCGGTGCGCTATACCGCTGCTCAAG	8870			
378	GGGCTCGCGCGCTTTCTGCACAAAGACATGCGCGGCGAGCTGTGTGGATGCGGTGGG	437			
8871	CGCGTGCAGCGCTGCTGAGAGTGTGCGGCGGTGAGAGTGTGCGGCGTGCATATCGCGCG	8930			
438	GGCGCGCGGAGTTTTCAGTGTGAGACCTCGGAGCCAGCTTCCAGTGCATGTGTACAAG	497			
8931	CGCGTGCAGAGGCGCTTCCGCGAGATGCTGCGGCAACGCGGCGCTGAGAGGCGAGGTGTG	8990			
498	GATGGCATGAGCTGGGCGCATCTCCGCTGAGCGCTTCTTTCAGAGGAGATGTGGGACGCGG	557			
8991	GTGAAACAAGGTCAAGGAGCTCTCCGCTGCGTTCGCTTCAACGCGCGCACCTAC	9050			
558	CACCGGCTGTGCGACGCACTACACAGGAGATGAGGCACTATCTCTGCGCGGTG	617			
9051	GAGGACTGCTGGCGCGCGCGTCTGACACCGCGCAAGGTGAGCGCGACGCGCTGCAG	9110			
618	GGCGAGGACTCTGTGAGCTTCCGCGTCCGCGTCTCTGAGCCCAAGTGTGTGTCTAAG	677			
9111	AACGCGGCTCGTGTCTGCTGATGTCTACACAGAGGCATGTGTGCGAGGCTCGG	9170			
678	GAGCAG 683				
9171	AAGGAG 9176				
RESULT 6					
US-09-620-312D-473					
; Sequence 473, Application US/09620312D					
; Patent No. 6569662					
; GENERAL INFORMATION:					
; APPLICANT: Tang, Y. Tom					
; APPLICANT: Liu, Chenghua					
; APPLICANT: Asundi, Vinod					
; APPLICANT: Zhang, Jie					
; APPLICANT: Ren, Feiyan					
; APPLICANT: Chen, Rui-hong					
; APPLICANT: Zhao, Qing A.					
; APPLICANT: Wehrman, Tom					
; APPLICANT: Xue, Aidong J.					
; APPLICANT: Yang, Yonghong					
; APPLICANT: Wang, Jian-Rui					
; APPLICANT: Zhou, Ping					
; APPLICANT: Ma, Yunding					
; APPLICANT: Wang, Dunrui					
; APPLICANT: Wang, Zhiwei					
; APPLICANT: John Tillinghast					
; APPLICANT: Drmanac, Radoje T.					
; TITLE OF INVENTION: NO. 6569662el Nucleic Acids and					
; TITLE OF INVENTION: Polypeptides					
; FILE REFERENCE: 784CIP2B					
; CURRENT APPLICATION NUMBER: US/09/620.312D					
; CURRENT FILING DATE: 2000-07-19					
; PRIOR APPLICATION NUMBER: 09/552,317					
; PRIOR FILING DATE: 2000-04-25					
; PRIOR APPLICATION NUMBER: 09/488,725					
; PRIOR FILING DATE: 2000-01-21					
; NUMBER OF SEQ ID NOS: 1105					


```
; SEQ ID NO 1
; LENGTH: 50937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Recombinant DNA
US-09-428-517-1

Query Match          4.9%; Score 49.4; DB 3; Length 50937;
Best Local Similarity 48.1%; Pred. No. 0.22;
Matches 140; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 76 GCATCGGATGAGTCCGGGCGGCCCCCAGAGGCTGACTGCCAAGCGCCGACGCGAG 135
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1146 GCAGACCATGAGGCGCTGTGAAACGCGGTGCCCATGTCGCGTGCAGATCGCGG 1205
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 136 CTGCCGCGAGTGTGCTCGCTGCGTGCACGAGGAGCGCAGTGTGCTGAGCTGTGCTAG 195
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1206 AGCAGACGATGAACCGCAGCGGATCGTGCAGCTGGGCTTCGGCCGGGCACATCCCGCGGG 1265
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 196 ATCAGGCTCGGCTGTGACGTGCTGAAAGGATGCTGCACACTGTGCCCGAGGCCCAAGT 255
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1266 ACCAGGTCAAGCGCAGAGAGCTGCGAGGCGGTGCTCGCGCTCGCCTCCGACCCCGGTG 1325
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 ATGAGGTGAGGATCGGCGGCGGCGGCGGTGCTCTCTGTGCGAGATGTGCGCGCGGAGC 315
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1326 TCGCCGAACGGCTCGCGGCGCTCGCGCAGAGATCCGTGAGCGCGGCGGCGCGCGCGG 1385
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 ATCAGGCTCTTACAGTGCCTGACGCGCGGCGGCGCATCGCCTACCGC 366
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1386 CGCCGACATCTGAGGCGATCTCGCCGAAGCAGGTGACCGCCCGCTGC 1436
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
US-09-540-8525
; Sequence 8525, Application US/0902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 8525
; LENGTH: 2670
; TYPE: DNA
; ORGANISM: Myxococcus xanthus
US-09-540-8525
```

```
Query Match          4.9%; Score 49; DB 4; Length 2670;
Best Local Similarity 43.8%; Pred. No. 0.15;
Matches 312; Conservative 0; Mismatches 395; Indels 6; Gaps 2;

QY 85 TGGAGGTCCGGGCGGCGCCCGGCTGACTGCCAAGCGCGCCAGCGCAGCTGCCCGGG 144
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1163 TGTGGGCGGCGCTGAGGCGGTGCGAAGGCGACGCGCGGGACTTCGCGACGTGCGGC 1222
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 145 AGGTGCTGGCTGCGACGAGGAGCGCAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 204
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1223 AGGTCCGGGAGCGGGTGTCTGCGCTGGGACTGTGTTGGCGGAGCTGAGGTGCGCGCG 1282
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 205 CGGCTGTGACGTGCTGAAGATGCTGCACACTGTCCCGAGGCCCAAGTATGAGGTGC 264
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1283 CGGCCGAGGACGCGGTGAGCGCGCGGCTCGTTCAACGGGGGCGCGCGCGCGAG 1342
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 265 AGGCATCGCGCGGCGGCGGTGCTCTCTGTGCGAGATGTGGCCCGGAGCGATGACGCC 324
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: March 21, 2005, 14:39:03
Job time : 188.678 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 20, 2005, 17:32:32 ; Search time 627.052 Seconds
(without alignments)
9503.485 Million cell updates/sec

Title: US-10-077-130-4_COPY_10500_11500

Perfect score: 1001
Sequence: 1 gccccttgagtgaggaa.....ttctagaagatgtgaaaaa 1001

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 5544816 seqs, 2976611598 residues

Total number of hits satisfying chosen parameters: 11089632

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

```
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1001	100.0	20489	18	US-10-723-860-5701 Sequence 5701, Ap
2	1001	100.0	23907	13	US-10-077-130-6 Sequence 6, Appli
3	1001	100.0	24120	13	US-10-077-130-4 Sequence 4, Appli
4	997.8	99.7	7564	18	US-10-476-397-15 Sequence 15, Appl
5	997.8	99.7	14061	17	US-10-093-463-73 Sequence 73, Appl
6	997.8	99.7	14109	17	US-10-093-463-71 Sequence 71, Appl
7	107.4	10.7	2534	17	US-10-108-260A-1946 Sequence 1946, Ap
8	62	6.2	3935	17	US-10-108-260A-534 Sequence 534, App
9	60	6.0	60	10	US-09-908-975-13516 Sequence 13516, A
10	59	5.9	82027	19	US-10-278-698-1034 Sequence 1034, Ap
11	59	5.9	82027	19	US-10-278-698-1045 Sequence 1045, Ap

12	57.4	5.7	81940	9	US-09-759-508B-1	Sequence 1, Appli
13	57.4	5.7	81940	10	US-09-960-706-1092	Sequence 1092, Ap
14	57.4	5.7	81940	10	US-09-873-319-724	Sequence 724, App
15	57.4	5.7	81940	18	US-10-723-860-132	Sequence 132, App
16	57.4	5.7	81940	18	US-10-656-873A-1	Sequence 1, Appli
17	53.6	5.4	2170	17	US-10-264-049-850	Sequence 850, App
18	52.6	5.3	1115	14	US-10-184-644-440	Sequence 440, App
19	52.6	5.3	1115	14	US-10-184-644-440	Sequence 440, App
20	52	5.2	1645	9	US-09-726-643-22	Sequence 22, Appl
21	52	5.2	1645	13	US-10-042-141-22	Sequence 22, Appl
22	52	5.2	1645	19	US-10-919-272-22	Sequence 22, Appl
23	51	5.1	2335	17	US-10-424-599-81504	Sequence 81504, A
24	49.8	5.0	2254	9	US-09-736-643-44	Sequence 44, Appl
25	49.8	5.0	2254	13	US-10-042-141-44	Sequence 44, Appl
26	49.8	5.0	2254	19	US-10-919-272-44	Sequence 117, App
27	47.6	4.8	716	10	US-09-822-846-117	Sequence 346, App
28	46.8	4.7	671	14	US-10-184-644-346	Sequence 346, App
29	46.8	4.7	671	14	US-10-184-644-346	Sequence 346, App
30	45.6	4.6	594	14	US-10-123-155-10	Sequence 10, Appl
31	45.6	4.6	594	15	US-10-146-731-10	Sequence 10, Appl
32	45.6	4.6	594	15	US-10-140-472-10	Sequence 10, Appl
33	45.6	4.6	594	15	US-10-141-761-10	Sequence 10, Appl
34	45.6	4.6	594	16	US-10-142-885-10	Sequence 10, Appl
35	45.6	4.6	594	16	US-10-158-790-10	Sequence 10, Appl
36	45.6	4.6	594	17	US-10-137-871-10	Sequence 10, Appl
37	45.6	4.6	594	17	US-10-140-923-10	Sequence 10, Appl
38	45.6	4.6	594	17	US-10-141-756-10	Sequence 10, Appl
39	45.6	4.6	594	17	US-10-141-759-10	Sequence 10, Appl
40	45.6	4.6	594	17	US-10-140-805-10	Sequence 10, Appl
41	45.6	4.6	594	17	US-10-140-864-10	Sequence 10, Appl
42	45.6	4.6	594	17	US-10-142-426-10	Sequence 10, Appl
43	44	4.4	1257	18	US-10-472-380-4	Sequence 1, Appli
44	44	4.4	1435	18	US-10-472-380-4	Sequence 4, Appli
45	43.8	4.4	882	14	US-10-184-644-574	Sequence 574, App

ALIGNMENTS

RESULT 1
US-10-723-860-5701
; Sequence 5701, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Nataasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods for Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723.860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5701
; LENGTH: 20489
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (565)..(584)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2938)..(2955)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (3329)..(3364)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:

NAME/KEY: misc_feature
LOCATION: (4556)..(4586)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4721)..(4735)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (4997)..(5011)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (5108)..(5138)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (6237)..(6300)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (7520)..(7566)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (12014)..(12037)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (12123)..(12146)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13258)..(13277)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (13918)..(13942)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15219)..(15248)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15262)..(15280)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15309)..(15381)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (15862)..(15882)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (17286)..(17303)
OTHER INFORMATION: n is a, c, g, or t
FEATURE:
NAME/KEY: misc_feature
LOCATION: (17413)..(17430)
OTHER INFORMATION: n is a, c, g, or t

Query Match 100.0%; Score 1001; DB 18; Length 20489;

Best Local Similarity 100.0%; Pred. No. 2.4e-268;
Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCCCTGTGGAGTGGAGAGGGGCCCGAGAACTCAGAGATGGGGACAGATACATCTG 60
Db 10500 GCCCTGTGGAGTGGAGAGGGGCCCGAGAACTCAGAGATGGGGACAGATACATCTG 10559

RESULT 2

US-10-077-130-6

; Sequence 6, Application US/10077130

; Publication No. US20020168742A1

; GENERAL INFORMATION:

; APPLICANT: Kapeller-Libermann, Rosana

; APPLICANT: Acton, Susan L.

; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family

Qy 61 AGGCAGAGGGGACCCAGGTGTGAGCTGCAGATCTGTGGCCTTGGCCATGGCGGACGCCGGG 120
Db 10560 AGGCAGAGGGGACCCAGGTGTGAGCTGCAGATCTGTGGCCTTGGCCATGGCGGACGCCGGG 10619
Qy 121 GAGTACTTGTGTGTGGGGCAGGAGAGACCTCAGCCAGCTCACCATCAGGGCTCTG 180
Db 10620 GAGTACTTGTGTGTGGGGCAGGAGAGACCTCAGCCACGCTCACCATCAGGGCTCTG 10679
Qy 181 CCTGCCAGGTTTCATAGAAGATGTGAAAAACCCAGAGGCCAGAGAGGGGCCACCGCTGTG 240
Db 10680 CCTGCCAGGTTTCATAGAAGATGTGAAAAACCCAGAGGCCAGAGAGGGGCCACCGCTGTG 10739
Qy 241 CTGCAGTGTGAGCTGAACAGTGCAGCCCTGTGTGAGTGGAGAAAGGGGTCTTGAGACCCCTC 300
Db 10740 CTGCAGTGTGAGCTGAACAGTGCAGCCCTGTGTGAGTGGAGAAAGGGGTCTTGAGACCCCTC 10799
Qy 301 AGAGATGGGACAGATACAGCTTGAGGCAGGACGGGACTAAATGTGAGCTGCAGATTCTG 360
Db 10800 AGAGATGGGACAGATACAGCTTGAGGCAGGACGGGACTAAATGTGAGCTGCAGATTCTG 10859
Qy 361 GGCCTGCCCATGGCAGACACTGGGGAGTACTCGTGGCTGTGCGGGCAGGAGAGACCTCG 420
Db 10860 GGCCTGCCCATGGCAGACACTGGGGAGTACTCGTGGCTGTGCGGGCAGGAGAGACCTCG 10919
Qy 421 GCTATGCTCAACCGTCAAGGCTCTACCCATCAAGTTTCACAGAGGGTCTTGAGAAACGAAGAG 480
Db 10920 GCTATGCTCAACCGTCAAGGCTCTACCCATCAAGTTTCACAGAGGGTCTTGAGAAACGAAGAG 10979
Qy 481 GCCACAGAGGGGCAACAGCCGTGCTGCGGTGTGAGCTGAGCAAGATGGCCCCCGTGGAG 540
Db 10980 GCCACAGAGGGGCAACAGCCGTGCTGCGGTGTGAGCTGAGCAAGATGGCCCCCGTGGAG 11039
Qy 541 TGGTGAAGGGGCATGAGACCTCAGAGATGGAGACAGACACAGCCCTGAGCGCAGGACGGG 600
Db 11040 TGGTGAAGGGGCATGAGACCTCAGAGATGGAGACAGACACAGCCCTGAGCGCAGGACGGG 11099
Qy 601 GCCAGTGTGAGCTGCAGATCCCGGCTCGTGGCAGAGACGCTGGGGAGTACCTGTGC 660
Db 11100 GCCAGTGTGAGCTGCAGATCCCGGCTCGTGGCAGAGACGCTGGGGAGTACCTGTGC 11159
Qy 661 ATGTGGGGAAGGAGAGGACCTCAGCCATGCTCACGTCAGGGCCATGCCCTTCAAGTTC 720
Db 11160 ATGTGGGGAAGGAGAGGACCTCAGCCATGCTCACGTCAGGGCCATGCCCTTCAAGTTC 11219
Qy 721 ATAGAGGCTCTGAGGAATGAAGAGGCCACAGAAAGGGACACGCGTGTGTGTGAG 780
Db 11220 ATAGAGGCTCTGAGGAATGAAGAGGCCACAGAAAGGGACACGCGCGTGTGTGTGAG 11279
Qy 781 CTGAGCAAGCGGCACCGGTGGAGTGGAGAGGGGCATGAGACCTCAGAGATGGGGAC 840
Db 11280 CTGAGCAAGCGGCACCGGTGGAGTGGAGAGGGGCATGAGACCTCAGAGATGGGGAC 11339
Qy 841 AGACACAGCCTGAGGACAGGACGGGTCCAGTGTGAGCTGCAGATCCGTGGCTGGCTGTG 900
Db 11340 AGACACAGCCTGAGGACAGGACGGGTCCAGTGTGAGCTGCAGATCCGTGGCTGGCTGTG 11399
Qy 901 GTGGATGCCGGGAGTACTCGTGTGTGGGGCAGGAGAGGACCTCAGCCACACTCACT 960
Db 11400 GTGGATGCCGGGAGTACTCGTGTGTGGGGCAGGAGAGGACCTCAGCCACACTCACT 11459
Qy 961 GTCAGGCGCCCTGCTGCCAGATTTCATAGAAGATGTGAAAA 1001
Db 11460 GTCAGGCGCCCTGCTGCCAGATTTCATAGAAGATGTGAAAA 11500

;; TITLE OF INVENTION: Members and Uses Therefor

;; FILE REFERENCE: MP12001-047P1RCP1(M)

;; CURRENT APPLICATION NUMBER: US/10/077,130

;; CURRENT FILING DATE: 2002-02-15

;; PRIOR APPLICATION NUMBER: 60/269201

;; PRIOR FILING DATE: 2001-02-15

;; NUMBER OF SEQ ID NOS: 9

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 6

;; LENGTH: 23907

;; TYPE: DNA

;; ORGANISM: Homo sapiens

US-10-077-130-6

Query Match 100.0%; Score 1001; DB 13; Length 23907;

Best Local Similarity 100.0%; Pred. No. 2.4e-268;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 GCCCTGTGGAGTGGAGGAGCGGCCGAGAACCTCAGAGATGGGACAGATACATCCTG 60
Db 10429 GCCCTGTGGAGTGGAGGAGCGGCCGAGAACCTCAGAGATGGGACAGATACATCCTG 10488

Qy 61 AGGCAGGAGGGACACAGGTGTGAGCTGCAGATCTGTGGCTTGGCCATGGCGGAGCGCGGG 120
Db 10489 AGGCAGGAGGGACACAGGTGTGAGCTGCAGATCTGTGGCTTGGCCATGGCGGAGCGCGGG 10548

Qy 121 GAGTACTTGTGTGTGCGGCGAGAGAGGACCTCAGCCACGCTCACCATCAGGGCTCTG 180
Db 10549 GAGTACTTGTGTGTGCGGCGAGAGAGGACCTCAGCCACGCTCACCATCAGGGCTCTG 10608

Qy 181 CCTGCCAGTTTCATAGAGATGTGAAACACAGAGGCCAGAGGGGCCACCGCTGTG 240
Db 10609 CCTGCCAGTTTCATAGAGATGTGAAACACAGAGGCCAGAGGGGCCACCGCTGTG 10668

Qy 241 CTGCAGTGTGAGCTGAAACAGTGCAGCCCTCTGAGTGGAGAAAGGGGTCTGAGACCTC 300
Db 10669 CTGCAGTGTGAGCTGAAACAGTGCAGCCCTCTGAGTGGAGAAAGGGGTCTGAGACCTC 10728

Qy 301 AGAGATGGGACAGATACAGCTTGAAGCAGGAGCGGACCTAAATGTAGCTGAGATTGCT 360
Db 10729 AGAGATGGGACAGATACAGCTTGAAGCAGGAGCGGACCTAAATGTAGCTGAGATTGCT 10788

Qy 361 GGCCTGGCCATGGCAGACACTGGGAGTACTCGTGGTGGCGGCGAGGAGGACCTCG 420
Db 10789 GGCCTGGCCATGGCAGACACTGGGAGTACTCGTGGTGGCGGCGAGGAGGACCTCG 10848

Qy 421 GCTATGCTCACCGTTCAGGGCTTACCCATCAAGTTTCAAGAGGGTCTGAGGAAAGAGAG 480
Db 10849 GCTATGCTCACCGTTCAGGGCTTACCCATCAAGTTTCAAGAGGGTCTGAGGAAAGAGAG 10908

Qy 481 GCCACAGAGGGGCAACAGCCGTGTGCGGTGTGAGCTGAGCAAGATGGCCCGCTGGAG 540
Db 10909 GCCACAGAGGGGCAACAGCCGTGTGCGGTGTGAGCTGAGCAAGATGGCCCGCTGGAG 10968

Qy 541 TGGTGAAGGGGCATGAGACCTCAGAGATGGAGACAGACACAGCCTGAGCAGGACGGG 600
Db 10969 TGGTGAAGGGGCATGAGACCTCAGAGATGGAGACAGACACAGCCTGAGCAGGACGGG 11028

Qy 601 GCCAGGTGTGAGCTGCAGATCCGGGCTCTGTCGAGAGGACGCTGGGGAGTACTGTGC 660
Db 11029 GCCAGGTGTGAGCTGCAGATCCGGGCTCTGTCGAGAGGACGCTGGGGAGTACTGTGC 11088

Qy 661 ATGTGCGGGAAGGAGGACCTCAGCCATGCTCACCGTCAGGGCCATGCCCTTCCAAGTTC 720
Db 11089 ATGTGCGGGAAGGAGGACCTCAGCCATGCTCACCGTCAGGGCCATGCCCTTCCAAGTTC 11148

Qy 721 ATAGAGGTCTGAGGAATGAAGAGCCACAGAGGGGACAGGCCACGCTGTGTGTGAG 780
Db 11149 ATAGAGGTCTGAGGAATGAAGAGCCACAGAGGGGACAGGCCACGCTGTGTGTGAG 11208

Qy 781 CTGAGCAAGCGGCAACCGGTGAGTGGAGGAGGGGCATGAGCCTCAGAGATGGGGAC 840
Db 11209 CTGAGCAAGCGGCAACCGGTGAGTGGAGGAGGGGCATGAGCCTCAGAGATGGGGAC 11268
```

RESULT 3

US-10-077-130-4

;; Sequence 4, Application US/10077130

;; Publication No. US20020168742A1

;; GENERAL INFORMATION:

;; APPLICANT: Kapeller-Libermann, Rosana

;; APPLICANT: Acton, Susan L.

;; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family

;; TITLE OF INVENTION: Members and Uses Therefor

;; FILE REFERENCE: MP12001-047P1RCP1(M)

;; CURRENT APPLICATION NUMBER: US/10/077,130

;; CURRENT FILING DATE: 2002-02-15

;; PRIOR APPLICATION NUMBER: 60/269201

;; PRIOR FILING DATE: 2001-02-15

;; NUMBER OF SEQ ID NOS: 9

;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 4

;; LENGTH: 24120

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: 5'UTR

;; LOCATION: (1)...(71)

;; NAME/KEY: CDS

;; LOCATION: (72)...(23978)

;; NAME/KEY: 3'UTR

;; LOCATION: (23979)...(24120)

US-10-077-130-4

Query Match 100.0%; Score 1001; DB 13; Length 24120;

Best Local Similarity 100.0%; Pred. No. 2.4e-268;

Matches 1001; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 GCCCTGTGGAGTGGAGGAGCGGCCGAGAACCTCAGAGATGGGACAGATACATCCTG 60
Db 10500 GCCCTGTGGAGTGGAGGAGCGGCCGAGAACCTCAGAGATGGGACAGATACATCCTG 10559

Qy 61 AGGCAGGAGGGACACAGGTGTGAGCTGCAGATCTGTGGCTTGGCCATGGCGGAGCGCGGG 120
Db 10560 AGGCAGGAGGGACACAGGTGTGAGCTGCAGATCTGTGGCTTGGCCATGGCGGAGCGCGGG 10619

Qy 121 GAGTACTTGTGTGTGCGGCGAGAGAGGACCTCAGCCACGCTCACCATCAGGGCTCTG 180
Db 10620 GAGTACTTGTGTGTGCGGCGAGAGAGGACCTCAGCCACGCTCACCATCAGGGCTCTG 10679

Qy 181 CCTGCCAGTTTCATAGAGATGTGAAACACAGAGGCCAGAGGGGCCACCGCTGTG 240
Db 10680 CCTGCCAGTTTCATAGAGATGTGAAACACAGAGGCCAGAGGGGCCACCGCTGTG 10739

Qy 241 CTGCAGTGTGAGCTGAAACAGTGCAGCCCTCTGAGTGGAGTGGAGAAAGGGGTCTGAGACCTC 300
Db 10740 CTGCAGTGTGAGCTGAAACAGTGCAGCCCTCTGAGTGGAGTGGAGAAAGGGGTCTGAGACCTC 10799

Qy 301 AGAGATGGGACAGATACAGCCTGAGCAGGACGCGGACTAAATGTAGCTGAGATTGCT 360
Db 10800 AGAGATGGGACAGATACAGCCTGAGCAGGACGCGGACTAAATGTAGCTGAGATTGCT 10859

Qy 361 GGCCTGGCCATGGCAGACACTGGGAGTACTCGTGGTGGCGGCGAGGAGGACCTCG 420
```


Db 10860 GGCCTGGCCATGGCAGACACTCGGGGAGTACTCGTGTGCGGTGTCGGGCGCAGGAGACCTCTG 10919
Qy 421 GCTATGCTCACCGTCAGGGCTCTACCCATCAAGTTTCAGAGGGTCTCAGGAACGAAGAG 480
Db 10920 GCTATGCTCACCGTCAGGGCTCTACCCATCAAGTTTCAGAGGGTCTCAGGAACGAAGAG 10979
Qy 481 GCCACAGAAGGGGCAACAGCGGTGTCGGGTGTGAGCTGTAGCAAGATGGCCCCCGTGGAG 540
Db 10980 GCCACAGAAGGGGCAACAGCGGTGTCGGGTGTGAGCTGTAGCAAGATGGCCCCCGTGGAG 11039
Qy 541 TGGTGGAGGGGCGCATGAGACCTCAGAGATGGAGACAGACACAGCCTGAGGACGCGG 600
Db 11040 TGGTGGAGGGGCGCATGAGACCTCAGAGATGGAGACAGACACAGCCTGAGGACGCGG 11099
Qy 601 GCCAGGTGTGAGCTGCAGATCCGGGCTCGTGGCAGAGGACGCTGGGGAGTACTGTGC 660
Db 11100 GCCAGGTGTGAGCTGCAGATCCGGGCTCGTGGCAGAGGACGCTGGGGAGTACTGTGC 11159
Qy 661 ATGTGCGGGAAGGAGAGGACCTCAGCCATGCTCACCGTCAGGGCATGCCCTTCCAAAGTTC 720
Db 11160 ATGTGCGGGAAGGAGAGGACCTCAGCCATGCTCACCGTCAGGGCATGCCCTTCCAAAGTTC 11219
Qy 721 ATAGAGGTCTGAGGATGAGAGGCCACAGAAAGGGGACAGCGCCAGCTGTGTGTGAG 780
Db 11220 ATAGAGGTCTGAGGATGAGAGGCCACAGAAAGGGGACAGCGCCAGCTGTGTGTGAG 11279
Qy 781 CTGAGCAAGGCGCACCGGTGGAGTGGAGGAAGGGGCATGAGACCTCAGAGATGGGGAC 840
Db 11280 CTGAGCAAGGCGCACCGGTGGAGTGGAGGAAGGGGCATGAGACCTCAGAGATGGGGAC 11339
Qy 841 AGACACAGCCTGAGGCAAGGACGGGTCCAGGTGTGAGCTGTAGATCCGTGGCTGGCTGTG 900
Db 11340 AGACACAGCCTGAGGCAAGGACGGGTCCAGGTGTGAGCTGTAGATCCGTGGCTGGCTGTG 11399
Qy 901 GTGATGCCGGGGAGTACTCTGTGTGTGTCGGGCGCAGAGAGACCTCAGCCACTACT 960
Db 11400 GTGATGCCGGGGAGTACTCTGTGTGTGTCGGGCGCAGAGAGACCTCAGCCACTACT 11459
Qy 961 GTCAGGGCCCTGCTGCCAGATTCATAGAAATGTGAAAA 1001
Db 11460 GTCAGGGCCCTGCTGCCAGATTCATAGAAATGTGAAAA 11500

RESULT 4

US-10-476-397-15

; Sequence 15, Application US/10476397

; Publication No. US20040115687A1

; GENERAL INFORMATION:

; APPLICANT: YUE, Henry

; APPLICANT: LEE, Ernestine A.

; APPLICANT: DUGGAN, Brendan M.

; APPLICANT: THANGAVELU, Kavitha

; APPLICANT: HONCHELL, Cynthia D.

; APPLICANT: DING, Li

; APPLICANT: JACKSON, Jennifer L.

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: KALLICK, Deborah A.

; APPLICANT: LEE, Sally

; APPLICANT: WARREN, Bridget A.

; APPLICANT: XU, Yuming

; APPLICANT: TRAN, Uyen K.

; APPLICANT: LAL, Preeti G.

; APPLICANT: THORNTON, Michael B.

; APPLICANT: HAFALIA, April J.A.

; APPLICANT: YAO, Monique G.

; APPLICANT: NGUYEN, Dannel B.

; APPLICANT: GANDHI, Ameena R.

; APPLICANT: KHAN, Farrah A.

; APPLICANT: CHAWLA, Narinder K.

; APPLICANT: GRIFFIN, Jennifer A.

; APPLICANT: CHINN, Anna M.

; APPLICANT: ELLIOTT, Vicki S.

; APPLICANT: RAMKUMAR, Jayalaxmi

; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: FORSYTHE, Ian J.
; TITLE OF INVENTION: CELL ADHESION AND EXTRACELLULAR MATRIX PROTEINS
; FILE REFERENCE: PF-0968 USN
; CURRENT APPLICATION NUMBER: US/10/476,397
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: PCT/US02/13874
; PRIOR FILING DATE: 2002-05-01
; PRIOR APPLICATION NUMBER: US 60/288,290
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/292,468
; PRIOR FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: US 60/298,616
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/301,672
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/345,008
; PRIOR FILING DATE: 2002-01-04
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 7564
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 7326129CB1
US-10-476-397-15

Query Match 99.7%; Score 997.8; DB 18; Length 7564;

Best Local Similarity 99.8%; Pred. No. 1.8e-267;

Matches 999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	GCCTCTGTGAGTGGAGGAGGGCCGAGAACTCAGAGATGGGACAGATACATCTG	60
Db	4546	GCCTCTGTGAGTGGAGGAGGGCCGAGAACTCAGAGATGGGACAGATACATCTG	4605
Qy	61	AGGCAGAGGGGACCCAGGTGTGAGCTGTGAGCTGTGGCTGTGGCCATGGCGGACGCCGGG	120
Db	4606	AGGCAGAGGGGACCCAGGTGTGAGCTGTGAGCTGTGGCTGTGGCCATGGCGGACGCCGGG	4665
Qy	121	GAGTACTTGTGTGTGCGGGCAGGAGAGACCTCAGCCAGCTCACCATCAGGGCTCTG	180
Db	4666	GAGTACTTGTGTGTGCGGGCAGGAGAGACCTCAGCCAGCTCACCATCAGGGCTCTG	4725
Qy	181	CCTGCCAGGTTTCATAGAAATGTGAAAAACAGAGAGCCAGAGAGGGGCCACGGCTGTG	240
Db	4726	CCTGCCAGGTTTCATAGAAATGTGAAAAACAGAGAGCCAGAGAGGGGCCACGGCTGTG	4785
Qy	241	CTGCAGTGTGAGCTGAACAGTGCAGGCCCTGTGTGGAGTGGAGAAAGGGGTCTGAGACCTC	300
Db	4786	CTGCAGTGTGAGCTGAACAGTGCAGGCCCTGTGTGGAGTGGAGAAAGGGGTCTGAGACCTC	4845
Qy	301	AGAGATGGGACAGATACAGCTTGGCAGAGACGGGACTAAATGTGAGCTGCAGATTCGT	360
Db	4846	AGAGATGGGACAGATACAGCTTGGCAGAGACGGGACTAAATGTGAGCTGCAGATTCGT	4905
Qy	361	GGCTGGCCATGCGAGACACTGGGGAGTACTCTGTGTGTGTGGGGCAGGAGGACCTCG	420
Db	4906	GGCTGGCCATGCGAGACACTGGGGAGTACTCTGTGTGTGTGGGGCAGGAGGACCTCG	4965
Qy	421	GCTATGCTCACCGTCAGGGCTCTACCCATCAAGTTTCAGAGGGTCTGAGGAACGAAGAG	480
Db	4966	GCTATGCTCACCGTCAGGGCTCTACCCATCAAGTTTCAGAGGGTCTGAGGAACGAAGAG	5025
Qy	481	GCCACAGAAGGGGCAACAGCGGTGTCGGGTGTGAGCTGTAGCAAGATGGCCCCCGTGGAG	540
Db	5026	GCCACAGAAGGGGCAACAGCGGTGTCGGGTGTGAGCTGTAGCAAGATGGCCCCCGTGGAG	5085
Qy	541	TGGTGGAAAGGGGCGCATGAGACCTCAGAGATGGAGACAGACACAGCCTGAGGACGCGG	600
Db	5086	TGGTGGAAAGGGGCGCATGAGACCTCAGAGATGGAGACAGACACAGCCTGAGGACGCGG	5145

Qy 601 GCCAGGTGTGAGCTGCAGATCCGGGCTCTGTGGCAGAGGACGCTGGGGAGTACTGTGC 660
Db |||||||
Qy 5146 GCCAGGTGTGAGCTGCAGATCCGGGCTCTGTGGCAGAGGACGCTGGGGAGTACTGTGC 5205
Db |||||||
Qy 661 ATGTGCGGGAAGGAGAGACCTCAGCCATGTCTCACCGTCAGGGCCATCCCTTCCAAAGTTC 720
Db |||||||
Qy 5206 ATGTGCGGGAAGGAGAGACCTCAGCCATGTCTCACCGTCAGGGCCATCCCTTCCAAAGTTC 5265
Db |||||||
Qy 721 ATAGAGGCTCTGAGGAATGAAGAGGCCACAGAGGGGACAGCGGCACGCTGTGTGTGAG 780
Db |||||||
Qy 5266 ATAGAGGCTCTGAGGAATGAAGAGGCCACAGAGGGGACAGCGGCACGCTGTGTGTGAG 5325
Db |||||||
Qy 781 CTGAGCAAGCGGCGCACCGGTGGAGTGGAGGAGGGGCATGAGCCTCAGAGATGGGGAC 840
Db |||||||
Qy 5326 CTGAGCAAGCGGCGCACCGGTGGAGTGGAGGAGGGGCATGAGCCTCAGAGATGGGGAC 5385
Db |||||||
Qy 841 AGACACAGCTTGAGGACGAGCGGTCCAGGTGTGAGCTGCAGATCCGTGGGCTGGCTGTG 900
Db |||||||
Qy 5386 AGACACAGCTTGAGGACGAGCGGTCCAGGTGTGAGCTGCAGATCCGTGGGCTGGCTGTG 5445
Db |||||||
Qy 901 GTGATGCGGGGAGTACTCTGTGTGTGGCGGACGAGAGACCTCAGCCACACTCACT 960
Db |||||||
Qy 5446 GTGATGCGGGGAGTACTCTGTGTGTGGCGGACGAGAGACCTCAGCCACACTCACT 5505
Db |||||||
Qy 961 GTGAGGCGCTGCTGCCAGATTCATAGAAGATGTGAAAA 1001
Db |||||||
Qy 5506 GTGAGGCGCTGCTGCCAGATTCATAGAAGATGTGAAAA 5546
Db |||||||

RESULT 5

US-10-093-463-73
; Sequence 73, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Ketha, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glennnda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: No. US20030208039A1el Antibodies that Bind to Antigenic Polypeptide
; TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.
; FILE REFERENCE: 21402-290A (Cura 590AT)
; CURRENT APPLICATION NUMBER: US/10/093,463
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283,675
; PRIOR FILING DATE: 2001-04-14
; PRIOR APPLICATION NUMBER: 60/338,092
; PRIOR FILING DATE: 2001-12-03

; PRIOR APPLICATION NUMBER: 60/274,281
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,101
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/325,681
; PRIOR FILING DATE: 2001-09-27
; PRIOR APPLICATION NUMBER: 60/304,354
; PRIOR FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/279,995
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 60/294,899
; PRIOR FILING DATE: 2001-05-31
; PRIOR APPLICATION NUMBER: 60/287,424
; PRIOR FILING DATE: 2001-04-30
; PRIOR APPLICATION NUMBER: 60/299,027
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/309,198
; PRIOR FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: 60/281,194
; PRIOR FILING DATE: 2001-04-04
; PRIOR APPLICATION NUMBER: 60/274,194
; PRIOR FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: 60/274,849
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: 60/330,380
; PRIOR FILING DATE: 2001-10-18
; PRIOR APPLICATION NUMBER: 60/275,235
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: 60/288,342
; PRIOR FILING DATE: 2001-05-03
; PRIOR APPLICATION NUMBER: 60/275,578
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 370
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 14061
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (15)..(14039)
US-10-093-463-73

Query Match 99.7%; Score 997.8; DB 17; Length 14061;
Best Local Similarity 99.8%; Pred. No. 1.8e-267;
Matches 999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GCCCTGTGGAGTGGAGAGGGGCGCGAGAACCTCAGAGATGGGACAGATACATCTCTG 60
Db |||||||
Qy 2661 GCCCTGTGGAGTGGAGAGGGGCGCGAGAACCTCAGAGATGGGACAGATACATCTCTG 2720
Db |||||||
Qy 61 AGGCAGGAGGGGAC CAGGTGTGAGCTGTGAGCTGTGGCCTTGGCCATGGCGGACGCCGGG 120
Db |||||||
Qy 2721 AGGCAGGAGGGGAC CAGGTGTGAGCTGTGAGCTGTGGCCTTGGCCATGGCGGACGCCGGG 2780
Db |||||||
Qy 121 GAGTACTTGTGTGTGGGCGCAGAGAGGACCTCAGCCAGCTCACCATCAGGGCTCTG 180
Db |||||||
Qy 2781 GAGTACTTGTGTGTGGGCGCAGAGAGGACCTCAGCCAGCTCACCATCAGGGCTCTG 2840
Db |||||||
Qy 181 CCTGCCAGGTTTCATAGAAGATGTGAAAAAACAGGAGGCGCAGAGAGGGGCGCACGCCCTGTG 240
Db |||||||
Qy 2841 CCTGCCAGGTTTCATAGAAGATGTGAAAAAACAGGAGGCGCAGAGAGGGGCGCACGCCCTGTG 2900
Db |||||||
Qy 241 CTGCAGTGTGAGCTGAACAGTGCAGCCCTCTGTGGAGTGGAGAAAGGGGTCTTGAGACCCCTC 300
Db |||||||
Qy 2901 CTGCAGTGTGAGCTGAACAGTGCAGCCCTCTGTGGAGTGGAGAAAGGGGTCTTGAGACCCCTT 2960
Db |||||||
Qy 301 AGAGATGGGACAGATACAGCTCAGGCGAGACCGGACTAAATGTGAGCTGCAGATTCTG 360
Db |||||||
Qy 2961 AGAGATGGGACAGATACAGCTCAGGCGAGACCGGACTAAATGTGAGCTGCAGATTCTG 3020
Db |||||||
Qy 361 GGCCTGGCCATGGCAGACACTGGGAGTACTCTGCTGGCTGGCGGACGAGGACCTCG 420
Db |||||||

Db 3021 GGCCTGGCCATGGCAGACACTGGGGAGTACTCGTGGTGTGCGGCGCAGGAGGACCTCG 3080
QY 421 GCTATGCTCACCGTCAAGGCTCTACCCATCAAGTTTCACAGAGGGTCTGAGGAACGAAGAG 480
Db 3081 GCTATGCTCACCGTCAAGGCTCTACCCATCAAGTTTCACAGAGGGTCTGAGGAACGAAGAG 3140
QY 481 GCCACAGAGGGGCAACAGCCGTGCTGCGGTGTGAGCTGAGCAAGATGGCCCCCTGGAG 540
Db 3141 GCCACAGAGGGGCAACAGCCGTGCTGCGGTGTGAGCTGAGCAAGATGGCCCCCTGGAG 3200
QY 541 TGGTGGAGGGGCATGAGACCTCTCAGAGATGAGAGCAGACACAGACCTGAGGACGAGCGG 600
Db 3201 TGGTGGAGGGGCATGAGACCTCTCAGAGATGAGAGCAGACACAGACCTGAGGACGAGCGG 3260
QY 601 GCCAGGTGTGAGCTGAGATCGCGGCTCTCGTGGCAGAGAGCGCTGGGGAGTACCTGTGC 660
Db 3261 GCCAGGTGTGAGCTGAGATCGCGGCTCTCGTGGCAGAGAGCGCTGGGGAGTACCTGTGC 3320
QY 661 ATGTGCGGAAGGAGGAGACCTCAGCCATGCTTCAACGTGAGGGCCATGCCCTTCCAAGTTC 720
Db 3321 ATGTGCGGAAGGAGGAGACCTCAGCCATGCTTCAACGTGAGGGCCATGCCCTTCCAAGTTC 3380
QY 721 ATAGAGGGTCTCAGGAATGAGAGGCCACAGAGGGGACACGGCCACGCTGTGGTGTGAG 780
Db 3381 ATAGAGGGTCTCAGGAATGAGAGGCCACAGAGGGGACACGGCCACGCTGTGGTGTGAG 3440
QY 781 CTGAGCAAGCGGCACCGGTGAGTGGAGAGGGGCGCATGAGACCTCAGAGATGGGGAC 840
Db 3441 CTGAGCAAGCGGCACCGGTGAGTGGAGAGGGGCGCATGAGACCTCAGAGATGGGGAC 3500
QY 841 AGACACAGCTGAGCAGAGACGGGTCCAGGTGTGAGCTGCGAGATCCGTGGGCTGGCTGTG 900
Db 3501 AGACACAGCTGAGCAGAGACGGGTCCAGGTGTGAGCTGCGAGATCCGTGGGCTGGCTGTG 3560
QY 901 GTGGATGCCGGGAGTACTCGTGTGTGTGTCGGGAGGAGAGACCTCAGCCACTCACT 960
Db 3561 GTGGATGCCGGGAGTACTCGTGTGTGTGTCGGGAGGAGAGACCTCAGCCACTCACT 3620
QY 961 GTCAGGCGCTCCCTGCCAGATTCATAGAAGATGTGAAAA 1001
Db 3621 GTCAGGCGCTCCCTGCCAGATTCATAGAAGATGTGAAAA 3661

RESULT 6

US-10-093-463-71
; Sequence 71, Application US/10093463
; Publication No. US20030208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Meres, Peter
; APPLICANT: Smithson, Glennda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zethusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corine
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward

Query Match 99.7%; Score 997.8; DB 17; Length 14109;
Best Local Similarity 99.8%; Pred. No. 1.8e-267;
Matches 999; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCCCTGTGGAGTGGAGGAGGGGCCGAGAACCTCAGAGATGGGACAGATACATCCTG 60
Db 2661 GCCCTGTGGAGTGGAGGAGGGGCCGAGAACCTCAGAGATGGGACAGATACATCCTG 2720
QY 61 AGGCAGGAGGGGACACAGGTGTGAGCTGAGATCTGTGGCTTGGCCATGGGACGCGGG 120
Db 2721 AGGCAGGAGGGGACACAGGTGTGAGCTGAGATCTGTGGCTTGGCCATGGGACGCGGG 2780
QY 121 GAGTACTTGTGTGTGTCGGGAGGAGAGACCTCAGCCACGCTCACCATCAGGGGCTCTG 180
Db 2781 GAGTACTTGTGTGTGTCGGGAGGAGAGACCTCAGCCACGCTCACCATCAGGGGCTCTG 2840

181 CCTGCCAGGTTTCATAGAGATGTGAAAAACAGGAGGCGCAGAGAAAGGGCCACCGCTGTG 240
Db |||||
2841 CCTGCCAGGTTTCATAGAGATGTGAAAAACAGGAGGCGCAGAGAAAGGGCCACCGCTGTG 2900
Qy |||||
241 CTCAGTGTGAGCTGAACAGTGCAGCCCTCTGAGAGTGGAGAAAGGGGTCTGAGACCTTC 300
Db |||||
2901 CTCAGTGTGAGCTGAACAGTGCAGCCCTCTGAGAGTGGAGAAAGGGGTCTGAGACCTTC 2960
Qy |||||
301 AGAGATGGGGACAGATACAGCTGAGGCGAGGACCGGACTTAATGTGAGCTGCAGATTTCGT 360
Db |||||
2961 AGAGATGGGGACAGATACAGCTGAGGCGAGGACCGGACTTAATGTGAGCTGCAGATTTCGT 3020
Qy |||||
361 GGCCTGGCCATGGGACAGACACTGGGAGTACTCTGCTGCTGCGGACAGGAGGACCTCG 420
Db |||||
3021 GGCCTGGCCATGGGACAGACACTGGGAGTACTCTGCTGCTGCGGACAGGAGGACCTCG 3080
Qy |||||
421 GCTATGCTCACCGTCCAGGCTCTACCCATCAAGTTTCACAGAGGGTCTGAGGAAACGAGAG 480
Db |||||
3081 GCTATGCTCACCGTCCAGGCTCTACCCATCAAGTTTCACAGAGGGTCTGAGGAAACGAGAG 3140
Qy |||||
481 GCCACAGAGGGGCAACAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db |||||
3141 GCCACAGAGGGGCAACAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3200
Qy |||||
541 TGCTGGAAGGGGCAACAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
Db |||||
3201 TGCTGGAAGGGGCAACAGCGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3260
Qy |||||
601 GCCAGGTGTGAGCTGCAGATCCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
Db |||||
3261 GCCAGGTGTGAGCTGCAGATCCGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3320
Qy |||||
661 ATGTGCGGGAAGGAGAGACCTCAGCCATGCTCACCCTCAGGCGCATGCCCTTCCAAAGTTC 720
Db |||||
3321 ATGTGCGGGAAGGAGAGACCTCAGCCATGCTCACCCTCAGGCGCATGCCCTTCCAAAGTTC 3380
Qy |||||
721 ATAGAGGTCTGAGGAATGAAGAGCCACAGAGGGGACAGGGGACAGCCCTCTGCTGCTGAG 780
Db |||||
3381 ATAGAGGTCTGAGGAATGAAGAGCCACAGAGGGGACAGGGGACAGCCCTCTGCTGCTGAG 3440
Qy |||||
781 CTGAGCAAGCGGCGCACCGGTGAGTGGAGGAGGGGCAATGAGACCTCAGAGATGGGGAC 840
Db |||||
3441 CTGAGCAAGCGGCGCACCGGTGAGTGGAGGAGGGGCAATGAGACCTCAGAGATGGGGAC 3500
Qy |||||
841 AGACACAGCTTGAAGCAGGAGCGGTTCAGGTGTGAGCTGAGCTGCAGATCCGCTGCTGCTG 900
Db |||||
3501 AGACACAGCTTGAAGCAGGATGGGTCCAGGTGTGAGCTGAGCTGCAGATCCGCTGCTGCTG 3560
Qy |||||
901 GTGATGCGGGGAGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db |||||
3561 GTGATGCGGGGAGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3620
Qy |||||
3621 GTGATGCGGGGAGTACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3680

RESULT 7

US-10-108-260A-1946
; Sequence 1946, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No. US20040005560A1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1946
; LENGTH: 2534
; TYPE: DNA
; ORGANISM: Homo sapiens

US-10-108-260A-1946

Query Match 10.7%; Score 107.4; DB 17; Length 2534;
Best Local Similarity 47.7%; Pred. No. 6.5e-20;
Matches 420; Conservative 0; Mismatches 446; Indels 15; Gaps 3;
Qy 136 TGGCGGAGGAGAGGACCTCAGCCAGCTCACCATCAGGGCTCTGCCCTGCCAGGTTTCATA 195
Db |||||
199 TCCCTGAAGGTGTGACCTCTGCCCGCTTGAGGTCCGAGTGAAGCCGGTGGTGTCTCTG 258
Qy |||||
196 GAAGATGTGAAAAACAGGAGGCGCAGAGAAAGGGGCGCAGGCTGTCTGCTGCAGTGTGAGCT- 254
Db |||||
259 AAGCGCTGGATGACCTGTCCGACAGAGAGCGCGCACCCCTGGCCCTCGAGTGTGAAGTC 318
Qy |||||
255 --GAAACAGTGCAGCCCTGTGAGAGTGAAGAAAGGGGTCTGAGACCTCAGAGATGGGAC 312
Db |||||
319 TCTGACCCGAGGCGCCATGTGTGCGCAAGATGGCGTGCAGCTGGGCCCCAGTGAC 378
Qy |||||
313 AGATACAGCCTGAGGAGGACGGGACTAAATGTGAGCTGCAGATTCTGGGCTGGCCATG 372
Db |||||
379 AAGTATGACTTCTGCAACAGCGCGGCGCTCGTGGTGCATGACGTGAGCCCT 438
Qy |||||
373 GCAGACACTGGGGAGTACTCTGCTGCTGTGCGGCGAGGAGAGACCTCGGCTATGCTACCC 432
Db |||||
439 GAAGACGCGGCTGTACACTGCTGCAAGTGGCTCCGAGGAGACCCGGGCGGGTCCGC 498
Qy |||||
433 GTGAGGCTCTTACCCATCAAGTTTCAGAGAGGTCTGAGGAAACGAGAGGCCACAGAGGG 492
Db |||||
499 GTGACGATCTGACGCTGGGCTCACCAGAGGCTGAAGACAATGGAGGTGCTGGAAGGG 558
Qy |||||
493 GCAACAGCGTGTGCGGTGAGCTGAGCAAGATGGCCCCGCTGGAGTGGTGGAGGGG 552
Db |||||
559 GAAAGCTGCAGCTTGTGAGTGGCTCTGCTCCACAGAGTGCACGACCCCGGCTATGTGG 618
Qy |||||
553 CATGAGACCTTCAGAGATGAGACAGACACAGCTGAGGCGAGGAG- - - - -GGGGCC 603
Db |||||
619 ACAGTGGTGGGAGACAGTGGGAGCTCCAGCGCTTCCAGGGCCACAGCTCAGGGCCGA 678
Qy |||||
604 AGGTGTGAGCTGCAGATCCGCGGCTCTGTCGACAGAGACGCTGGGGAGTACCTGTGCAATG 663
Db |||||
679 AATACATCTCTGCTGGTCCGGGAGGCTGCACCAAGTGTATGCGGGGAGGTGGTCTTCTCT 738
Qy |||||
664 TGGGGAGGAGGAGGACCTCAGCCATGCTCACCCTCAGGGCCATGCCCTTCCAGTTCATA 723
Db |||||
739 GTGCGGGGCTTCACCTCCAGGCTCCTCATTTCTCAGAGAGAGGCGCGGCCACCATCATC 798
Qy |||||
724 GAGGCTCTGAGGAATGAAGAGGCCACAGAGGGGACACGGCCACGCTGTGTGTGAGCTG 783
Db |||||
799 AAGCCCCCTGGAGACAGTGGTGGCCGACGGGAGGACGCTGGAGTGGTGGTGGTGGTGG 858
Qy |||||
784 AGCAAGGCGG- - -CACCCTGGAGTGGAGGAGGGGCAATGAGACCTCAGAGATGGGGAC 840
Db |||||
859 TCAGGGCGGGGACCGCCGCTGCACTGGCTGAAGACAGAGGAGGCGCATCCGCAAGAGCCAG 918
Qy |||||
841 AGACACAGCCTGAGGAGGAGCGGTCCAGGTGTGAGCTGCAGATCCGCTGGGCTGGCTGTG 900
Db |||||
919 AAGTATGATGTGTCTGCGAGGGGACGATGCTGCCCATGCTGTGTGCTATCCGCGGGGCTCTG 978
Qy |||||
901 GTGGATCCCGGGAGTACTCTGCTGCTGCTGCGGGCAGGAGGAGGACCTCAGCCACACTCACT 960
Db |||||
979 AAGGACGGGCGGAGTACAGCTGTGAGGTGGAGGCTTCCAGAGACACAGCCAGGCTCCAT 1038
Qy |||||
961 GTGAGGCGGCTGCCCTGCCAGATTTCATAGAAGATGTGAAAAA 1001
Db |||||
1039 GTGGAAGAAAAGCAAACTGCTTCACAGAGGAGCTGACCAA 1079

RESULT 8

US-10-108-260A-534
; Sequence 534, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE


```
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108.260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 534
; LENGTH: 3935
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-108-260A-534

Query Match      6.2%; Score 62; DB 17; Length 3935;
Best Local Similarity 46.0%; Pred. No. 2.8e-07;
Matches 295; Conservative 0; Mismatches 335; Indels 12; Gaps 2;

Qy 309 GGACAGATACAGCTGAGCGGAGGAGCGGAGCTAAATGTGAGCTGAGATTCGAGATTCGTCGCTGGC 369
Db 1033 GGGCAAGTACGATGTGAAGCAGATGGGCACCAAGTACATGCTGCTTATTAGCAACGTGAA 1092
Qy 369 CATGGCAGACATCGGGAGTACTCGTGGTGTGGCGCAGGAGGACCTCGGCTATGCT 428
Db 1093 CATGAACGATGCTGGCATCTACGCCCTGTCGCTGGCGATGAGCGATGAGTCAGAGCT 1152
Qy 429 CACCGTCAGGGCTCTACCCATCAAGTTTCAAGAGGCTCTGAGGAACGAAGGGCCACAGA 488
Db 1153 CACAGTGTGATGAGCCACTGAATTTCTTGGGAGAGATGAAGCTGTGAAGGTGACAGA 1212
Qy 489 AGGGGCAACAGCGCTGCTGGGTGTGAGCTGAGCAAGATGGCCCCCGTGGAGTGGTGAA 548
Db 1213 GCGCCAGACACCTGTGTTTGAGATCCGCTCTCAAGAAAGAGGCCCACTTTGTGTGAA 1272
Qy 549 GGGGCATGAGACCTCAGAGATGAGCAGACA-----CAGCCTGAGGCGAGGACGG 599
Db 1273 GTTCAATGGGAAGAGCTGAAGAGGATGACAAGTATGAATCAAGTGTCCGAAGATGG 1332
Qy 600 GGCAGGTGTGAGCTGCAGATCCGCGCTCTGTCGACAGGAGCGCTGGGAGTACCTGTG 659
Db 1333 TCTGACGACACAGCTTAAGATTAGGATGCCAGACTCAGTGCAGCGCGAGTCTCTGC 1392
Qy 660 CATGTGGGGAAGAGAGGACCTCAGCATCTCAGCGTCAAGGCGCATGCTTCCAAGTT 719
Db 1393 TGAGGCGGGGAACCTGTGTACAAAAGGCGCACCTCACTATTGACCGCATCCCCATCAAGTT 1452
Qy 720 CATAGAGGCTCTGAGGAATGAAGAGGCCACAGAAGGGGACACGCGCTGTGGTGTGA 779
Db 1453 TGTGAGCAACCTCAAAATGTACGTGTGAAGAGAGGAGTCCGCGATGCCCTGGAGTGA 1512
Qy 780 GCTGA---GCAAGGCGGCACCGGTGGAGTGGAGGAAGGGGCATGAGACCTTCAGAGATGG 836
Db 1513 GCTGACATCCAAGGATGTGACACTCGCTGGAAGAGGATGGCGAGCTGTGATGCATGG 1572
Qy 837 GGACAGACACAGCTTGAAGGAGGAGCGGTCCAGGTGTGAGCTGAGATCCGTCGCTGGC 896
Db 1573 CACTAAGTACAGCATGAACCAAGAGGCAAGCAAGCAGAGCTGATCATCGAGGATGACA 1632
Qy 897 TGTGTGATCCCGGGAGTACTCGTGTGTGTCGGGGCAGGA 938
Db 1633 GCTCAGTATGCTGGCGAGTACATGTGTGGCCATGCAGGA 1674

RESULT 9
US-09-908-975-13516
; Sequence 13516, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: WASSERMAN, Alon
; APPLICANT: SHOSHAN, Avi
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Eli
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
```

```
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908.975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 13516
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-908-975-13516

Query Match      6.0%; Score 60; DB 10; Length 60;
Best Local Similarity 100.0%; Pred. No. 7.9e-07;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 703 GCCATGCTTCCAAAGTTTCATAGAGGGTCTCAGGAATGAAGAGGCCACAGAAAGGGGACAG 762
Db 1 GCCATGCTTCCAAAGTTTCATAGAGGGTCTCAGGAATGAAGAGGCCACAGAAAGGGGACAG 60

RESULT 10
US-10-278-698-1034
; Sequence 1034, Application US/10278698
; Publication No. US20050037344A1
; GENERAL INFORMATION:
; APPLICANT: PathoArray GmbH
; APPLICANT: Stuhlmuller, Bruno
; APPLICANT: Haupl, Thomas
; TITLE OF INVENTION: Nucleic Acid Array
; FILE REFERENCE: 030027US
; CURRENT APPLICATION NUMBER: US/10/278.698
; CURRENT FILING DATE: 2002-10-23
; NUMBER OF SEQ ID NOS: 1050
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1034
; LENGTH: 82027
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-278-698-1034

Query Match      5.9%; Score 59; DB 19; Length 82027;
Best Local Similarity 43.4%; Pred. No. 2.3e-06;
Matches 376; Conservative 0; Mismatches 485; Indels 6; Gaps 2;

Qy 141 GCAGGAGAGGACCTCAGCCACGCTCACCATCAGGGCTCTCGCTCCAGGTTTCATAGAAGA 200
Db 15676 GGAGCTGAGCTCCACAGCAGCAGCTGAAAGGTCTTAGAGGCCCGATCCCTACTTCACTGTGAA 15735
Qy 201 TGTGAATAAACCGAGGAGGCCAGAGAGGGCCACGGCTGTGCTGCAGTGTGAGCTGAACAG 260
Db 15736 ATTACATGACAAACTGCAGTGGAGAAAGATGAGATTACTTTGAAGTGTGAGTGAAGCA 15795
Qy 261 TGCAGCCCTCTGTGAGTGGAGAAAGGGGTCTGAGACCTTCAGAGATGGGACAGATACAG 320
Db 15796 AGATGTACCAAGTGAATGGTTCAAAGATGGTGAAGAGATTGTCTCCCTTACCCTAATATTC 15855
Qy 321 CCTGAGGACAGACGGGACTAAATGTGAGCTGCAGATTCGTGGCCCTGGCCATGGCAGACAC 380
Db 15856 TATCAAGGCAGATGGCTCGCCGCGCATCTTAAATAATCAAAAAGCGGACCTTAAAGATAA 15915
Qy 381 TGGGGAGTACTCGTGGCTGTGCGGCGAGGAGGACCTCGCTATGTCTACCGTCAGGGC 440
Db 15916 AGCGAATATGTGTGACTGTGSCACAGACCAAGCCAAATGTTACTTGTGAGGC 15975
Qy 441 TCTACCCATCAAGTTTCAGAGGGTCTGAGSAACGAAGAGGCCACAGAAAGGGGCAACAGC 500
Db 15976 TCGACTAATAAAGCTGAAAGCCCTCTGTACGGAGTAGAGGTGTTTGTGGTGAACAGC 16035
Qy 501 CGTCTCGGTGTGAGCTGAGCAAGATGGCCCCCGCTGGAGTGGTGGAGGGGCATG---A 557
```


Db 16036 CCACTTTGAAATTTGAACTTTCTGAACTGATGTTTCAAGCCAGTGGAGCTGAAGGACA 16095
Qy 558 GACCCTCAGAGATGGAGACACACAGCCTCAGGACGACCGGGCCAGGCTGTGAGCTGCA 617
Db 16096 GCGTTTGACAGCTTCCCTCTGACTGTGAAATCATTTGAGGATGGAAGAGCATATTTCTGAT 16155
Qy 618 GATCCGGGCTCTGTGGCAGAGGACGCTGGGGAGTACCTGTGTCATGTGCGGGAAGGAGAG 677
Db 16156 CTTTCATAAATGTCAGCTGGGTATGACAGGAGAGGTTTCCCTTCCAGGCTGCTAATGCCAA 16215
Qy 678 GACCTCAGCATGCTCACCCTCAGGCGCATCCCTTCCAAGTTTCATAGAGGCTGTGAGAA 737
Db 16216 ATCTGACAGCAATCTGAAAGTGAAGAAATTCCTTCTATCTTCATCACACCTCTCAGTGA 16275
Qy 738 TGAAGAGGCCACAGAAGGGGACACGGCCACGCTGTGTGTGAGCTGAGCAAGGGGCGC--- 794
Db 16276 TGTTAAAGCTTTCGAGAAAGATGAGGCTAAGTTTGAAGTGTGAAGTATCCAGGAGGCCAA 16335
Qy 795 ACCGGTGGAGTGGAGGAAGGGGATGAGACCTCAGAGATGGGACACACAGCCTGAG 854
Db 16336 AACATTCGTTGGCTTAAAGGAACCCAGGAAATCACAGGTGATGACAGATTTGAGCTTAT 16395
Qy 855 GCAGAGGGTCCAGGTGTGAGCTGCAGATCCGTGGCCTGTGTGTGGATGCCGGGA 914
Db 16396 AAGGATGGCACTAAGCAATTCATTTGGTGTCAAGTCAAGTCAAGTGGCAATGTATCATTTGAAGATGAAGCAAA 16455
Qy 915 GTACTGTGTGTGTGGGAGGAGAGACCTCAGCCACACTCACTGTGAGGGCCCTGCC 974
Db 16456 ATACATGTTTGAAGCTGAAGTATGACACAAAGTGGCAATGTATCATTTGAAGATGAAGCAAA 16515
Qy 975 TGCAGATTCATAGAAGATGTGAAAAA 1001
Db 16516 GCTCAAAATTCCTCACCCCTCTCAAAGA 16542

RESULT 11

US-10-278-698-1045
; Sequence 1045, Application US/10278698
; Publication No. US20050037344A1
; GENERAL INFORMATION:
; APPLICANT: Pathoarray GmbH
; APPLICANT: Stuhlmüller, Bruno
; APPLICANT: Haupl, Thomas
; FILE OF INVENTION: Nucleic Acid Array
; FILE REFERENCE: O30027US
; CURRENT APPLICATION NUMBER: US/10/278,698
; CURRENT FILING DATE: 2002-10-23
; NUMBER OF SEQ ID NOS: 1050
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1045
; LENGTH: 82027
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-278-698-1045

Query Match 5.9%; Score 59; DB 19; Length 82027;
Best Local Similarity 43.4%; Pred. No. 2.3e-06;
Matches 376; Conservative 0; Mismatches 485; Indels 6; Gaps 2;
Qy 141 CGAGAGAGGACCTCAGCCAGCTCACCACCTCAGGGCTCTGCTGCGAGGTTTCATAGAAGA 200
Db 15676 GGAGCTGAGCTCCACAGCACAGCTGAAGTCTTAGAGGCCGATCCCTACTTCACTGTGAA 15735
Qy 201 TGTGAAAACAGGAGCCACAGAGGGCCACCGCTGTGTGCTGAGTGTGAGCTGAACAG 260
Db 15736 ATTACATGACAAAACCTGCAGTGGAGAGGATGAGATTACTTTGAAGTGTGAAGTGAGCAA 15795
Qy 261 TGCAGCCCTCTGGAGTGGAGAAAGGGGTCTGAGACCCCTCAGAGATGGGACAGATACAG 320
Db 15796 AGATGTACCACTGAATGGTTCAAGATGGTGAAGAGATTGTCCCTTCAACCAATATTC 15855
Qy 321 CCTGAGGAGGAGCGGACTAAATGTGAGCTGCAGATTCGTGGCCTGGCCATGGCAGACAC 380

Query Match

5.7%; Score 57.4; DB 9; Length 81940;

Db 15856 TATCAAGCGCAGATGGCCTGCGCCGATCTTTAAAAATCAAAAAGCGGACCTTAAAGATAA 15915
Qy 381 TGGGGAGTACTCGTGTGCGGTGTCGGGAGGAGGACCTCGGCTATGCTCACCGTCAGGGC 440
Db 15916 AGGCGAATATGTGTGACTGTGACACAGACAGCAAGGCAATGTTACTGTGTAGGC 15975
Qy 441 TCTACCCATCAAGTTTCAGAGGGTCTGAGGAAAGAGAGGCCACAGAAAGGGGCAACAGC 500
Db 15976 TCGACTAATAAAGTGGAAAGCCCTCTGTCGGAGTAGAGGTGTTGTTGTGTGAAACAGC 16035
Qy 501 CGTCTCGGTGTGAGCTGAGCAAGATGGCCCGTGGAGTGGTGGAGGGGCGATG---A 557
Db 16036 CCACTTTGAAATTTGAACTTTCTGAACCTGATGTTTCAAGCCAGTGGAAAGCTGAAAGGACA 16095
Qy 558 GACCCTCAGAGATGGAGACACACAGCCTGAGGACGACCGGGCCAGGCTGTGAGCTGCA 617
Db 16096 GCGTTTGACAGCTTCCCTCTGACTGTGAATCATTTGAGGATGGAAGAGCATATTTCTGAT 16155
Qy 618 GATCCGGGCTCTGTGGCAGAGGACGCTGGGGAGTACCTGTGTCATGTGCGGGAAGGAGAG 677
Db 16156 CTTTCATAAATGTCAGCTGGGTATGACAGGAGAGGTTTCCCTTCCAGGCTGCTAATGCCAA 16215
Qy 678 GACCTCAGCATGCTCACCCTCAGGCGCATCCCTTCCAAGTTTCATAGAGGCTGTGAGAA 737
Db 16216 ATCTGACAGCAATCTGAAAGTGAAGAAATTCCTTCTATCTTCATCACACCTCTCAGTGA 16275
Qy 738 TGAAGAGGCCACAGAAGGGGACACGGCCACGCTGTGTGTGAGCTGAGCAAGGGGCGC--- 794
Db 16276 TGTTAAAGCTTTCGAGAAAGATGAGGCTAAGTTTGAAGTGTGAGTATCCAGGAGCCCA 16335
Qy 795 ACCGGTGGAGTGGAGGAAGGGGATGAGACCTCAGAGATGGGACACAGACAGCCTGAG 854
Db 16336 AACATTCGTTGGCTTAAAGGAACCCAGGAAATCACAGGTGATGACAGATTTGAGCTTAT 16395
Qy 855 GCAGAGGGTCCAGGTGTGAGCTGCAGATCCGTGGCCTGTGTGTGAGTGGATGCCGGGA 914
Db 16396 AAGGATGGCACTAAGCAATTCATTTGGTGTCAAGTCAAGTCAAGTGGCAATGTATCATTTGAAGATGAAGCAAA 16455
Qy 915 GTACTGTGTGTGTGGGAGGAGAGACCTCAGCCACACTCACTGTGAGGGCCCTGCC 974
Db 16456 ATACATGTTTGAAGCTGAAGTATGACACAAAGTGGCAATGTATCATTTGAAGATGAAGCAAA 16515
Qy 975 TGCAGATTCATAGAAGATGTGAAAAA 1001
Db 16516 GCTCAAAATTCCTCACCCCTCTCAAAGA 16542

RESULT 12

US-09-759-508B-1
; Sequence 1, Application US/09759508B
; Publication No. US20020182599A1
; GENERAL INFORMATION:
; APPLICANT: Fishman, Mark C.
; TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease
; FILE REFERENCE: 00786/381002
; CURRENT APPLICATION NUMBER: US/09/759,508B
; CURRENT FILING DATE: 2001-01-12
; PRIOR APPLICATION NUMBER: US 60/175,787
; PRIOR FILING DATE: 2000-01-12
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 81940
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (133)..(80910)
; OTHER INFORMATION:
US-09-759-508B-1

Beat	Local Similarity	43.3%	Pred. No. 6.3e-06;	Matches 375;	Conservative 0;	Mismatches 486;	Indels 6;	Gaps 2;
QY	141	GCAGGAGGACCTCAGCCACGCTCACCATCAGGGCTCTGCCTGCCAGGTTTCATGAAGA	200					
Db	15585	GGAGCTGAGCTCCACAGACACAGCTGAAGTCTTTAGAGGCCGATCCCTACTCTCACTGTGA	15644					
QY	201	TGTCAAAAACGAGAGGCCAGAGAGAGGGGCCAGCGTGTGCTCATGTGTAGCTGAACAG	260					
Db	15645	ATTACATGACAAAACTCGCTGGAGAAGGATGAGATTACTTTGAAGTGTGAAGTGAAGCA	15704					
QY	261	TGCAGCCCTGTGGAGTGGAAAAAGGGTCTGAGACCTCAGAGATGGGGACAGATACAG	320					
Db	15705	AGATGTACCACTGAAATGGTTCAAAGATGCTGAAGAGATTGTCCTTACCCCAATATTTC	15764					
QY	321	CCTGAGGACGAGCGGGACTAAATGTGAGCTGCAGATTTCGTGCGCTGGCCATGCGACAC	380					
Db	15765	TATCAAGGCAGATGGCTGCGCGCATCTTTAAAAATCAAAAAGGCGGACTTTAAAGATAA	15824					
QY	381	TGGGGAGTACTCTGTCGTGTGCGGGCAGAGAGGACTCGGCTATGCTCACGTCAGGGC	440					
Db	15825	AGGCGAATATGTGTGACTGTGTGGCAGACAACAGCAAGGCMAATGTGTACTGTGTGGCG	15884					
QY	441	TCTACCATTCAAGTTTCACAGAGGCTGTGAGAACGAAAGAGGCCACAGAAAGGGGCAACAGC	500					
Db	15885	TCGACTAATAGAGTGGAAAAAGCCTCTGTACGAGTAGAGGTGTTGTGTGTAACAGC	15944					
QY	501	CGTGCTCGGTGTGAGCTGAGCAAGATTGGCCCCGTGGAGTGGTGGAAAGGGCATG--A	557					
Db	15945	CCAATTGAAATTGAACTTTCTGAACTGATGTTACGCGCCAGTGAAGCTGAAAGGACA	16004					
QY	558	GACCCTCAGAGATGGAGACAGACACGCTGAGGACGAGGACGGGCCAGGTGTGACTGCA	617					
Db	16005	GCCTTTGACAGCTTCCCTGACTGTGAAATCATTTAGGATGAAAGACATATTCTGAT	16064					
QY	618	GATCCGCGGCTCGTGCCAGAGGACGCTGGGGAGTACCTGTGCATGTGCGGGAAGGAG	677					
Db	16065	CCTTCATAACTGTGCTGGGTATGACAGAGAGGTTTCTTCCAGGCTGCTTAATGCCAA	16124					
QY	678	GACCTCAGCCATGCTCACCGTCAGGGCCATGCTTCCAAGTTTCATAGAGGCTGTGAGAA	737					
Db	16125	ATCTCAGGCCAAATCTGAAAGTGAAGAAATGCGCTCTTATCTTCATCACACCTCTCAGTGA	16184					
QY	738	TGAAGAGGCCACAGAAGGGGACACGCGCCACGCTGTGGTGTGAGCTGAGCAAGGCGGC--	794					
Db	16185	TGTTAAAGTCTTCGAGAAAGATGAGGCTAAGTTTGAGTGTGAAGTATCCAGGAGGCCAA	16244					
QY	795	ACCGTGTGAGTGGAGGAAGGGCATGAGACCTCAGAGATGGGACAGACACAGCCTGAG	854					
Db	16245	AACATTCGTTGGCTTAAAAAGAACCCAGGAAATCAAGGTGATGACAGATTGAGCTTAT	16304					
QY	855	GCAGGACGGTCCAGGTGTGAGCTGCAGATCCGTGGCTGGCTGTGGTGTGGATGCCGGGA	914					
Db	16305	AAAGGATGGCACTTAAGCATTCATGTGTGATCAAGTCAGCTGCTTTTGAAGATGAAGCAAA	16364					
QY	915	GTACTCGTGTGTGCGGGCAGGAGAGGACCTCAGCCACACTTCACCTGTGCGGGCCTGCC	974					
Db	16365	ATACATGTTTGAAGCTGGAAGATAAGCACAAAGTGGCAAACTGATCATTTGAAGGATCCG	16424					
QY	975	TGCCAGATTTCATAGAAGATGTGAAAAA	1001					
Db	16425	GCTCAAATTCCTCACCCCTCTCAAAGA	16451					

RESULT 13
US-09-960-706-1092
; Sequence 1092, Application US/09960706
; Publication No. US20030134280A1
; GENERAL INFORMATION:
; APPLICANT: Munger, William E.
; TITLE OF INVENTION: Identifying Drugs for and Diagnosis of Benign Prostatic Hyperplasia
; TITLE OF INVENTION: Gene Expression Profiles
; FILE REFERENCES: 44921-5029-01US


```

Qy 441 TCTACCCATCAAGTTCACAGAGGCTCTGAGGAACGAGAGGCCACAGAGGGGCAACAGC 500
Db 15885 TCGACTAATAGAAGTGAAGAGCCTCTGTACGGAGTAGAGGTGTTTGTGTGAAACAGC 15944
Qy 501 CGTGTGGGTGTGAGCTGAGCAAGATGGCCCCGTGGAGTGTGTGAAGGGGCATG---A 557
Db 15945 CCACTTTGAATTTGAACCTTCTGAACTGATGTTTACGGCCAGTGGGAAGCTGGAAGGACA 16004
Qy 558 GACCTCAGAGATGGAGACAGACACAGCCTGAGGACAGACGGGGCCAGGTGTGAGTGCAC 617
Db 16005 GCCTTTGACAGCTTCCCTGCTGCTGAAATCATTTGAGGATGGAAAGAACATATTCTGAT 16064
Qy 618 GATCCGGGGCTCTGTCAGAGGAGCGTGGGGATCCTGTGCATGTGCGGAGGAGAG 677
Db 16065 CCTTCATAACTGTGAGTGGGTATGACAGGAGAGGTTTCCCTCCAGGCTGCTAATGCCAA 16124
Qy 678 GACCTCAGCCATGCTCACCGTCAGGGCCATGCCTTCCAAGTTTCATAGAGGTTCTGAGGAA 737
Db 16125 ATCTGACCCCAATCTGAAAGTGAAGAATTGCTCTTTATCTTCATCACACCTCTCAGTGA 16184
Qy 738 TGAAGAGCCACAGAGGGGACACGGCCACGCTGTGTGTGAGCTGAGCAAGGGCGC--- 794
Db 16185 TGTAAAGTCTTCGAGAAAGATGAGGCTTAAGTTTGAAGTGTGAAGTATCCAGGGAGCCCAA 16244
Qy 795 ACCGTTGGAGTGGAGGAAGGGGCATGAGACCTCAGAGATGGGACACAGACAGCCTGAG 854
Db 16245 AACATTCGGTTGGCTAAAGGAACCCAGGAATCACAGGTGATGACAGATTTGAGCTTAT 16304
Qy 855 GCAGGAGGGTCCAGGTGTGAGTGCAGATCCGTGGCTGGCTGTGTGGATGCCGGGA 914
Db 16305 AAAGGATGGCACTAAGCATTCAATGGTGATCAAGTCAGCTGCTTTTGAAGATGAAGCAA 16364
Qy 915 GTACTCGTGTGTGCGGGCAGGAGGACCTCAGCCACACTCACTGTTCAGGGCCCTGCC 974
Db 16365 ATACATGTTTGAAGCTGAAGATAGACACCAAGTGGCAAACTGATCATTTGAAGGAATCCG 16424
Qy 975 TGCCAGATTTCATAGAAGATGTGAAGAAA 1001
Db 16425 GCTCAATTCCTCACCCTCTCAAGA 16451

```

Search completed: March 21, 2005, 16:07:39
 Job time : 633.052 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 17:29:58 ; Search time 496 Seconds
(without alignments)
5309.984 Million cell updates/sec

Title: US-10-077-130-5

Perfect score: 41273

Sequence: 1 MDQPFSGARFLTRPKAFV.....NRREKRRALLYKRNLAQVR 7968

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1401741 seqs, 330541175 residues

Total number of hits satisfying chosen parameters: 1401741

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
17: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	41273	100.0	7968	13	US-10-077-130-5
2	32134	77.9	6620	15	US-10-080-334-2290
3	32134	77.9	6620	16	US-10-408-765A-2291
4	13710	33.2	2630	13	US-10-077-130-2
5	13528	32.8	2596	14	US-10-307-019-6
6	10519.5	25.5	4691	15	US-10-093-463-72
7	10506	25.5	4675	15	US-10-093-463-74
8	9834.5	23.8	2328	16	US-10-476-397-4
9	8423	20.4	1665	9	US-09-858-664A-2
10	8423	20.4	1665	14	US-10-274-978-2
11	8423	20.4	1665	15	US-10-697-263-2
12	8423	20.4	1665	15	US-10-921-168-2
13	8410	20.4	1665	15	US-10-415-011-22

14	8407	20.4	1618	15	US-10-182-243-46	Sequence 46, Appl
15	8403	20.4	1610	14	US-10-307-019-4	Sequence 4, Appl
16	7073	17.1	1351	14	US-10-307-019-1	Sequence 1, Appl
17	6619.5	16.0	1596	16	US-10-408-765A-992	Sequence 992, App
18	4533	11.0	871	14	US-10-307-019-7	Sequence 7, Appl
19	4529	11.0	871	15	US-10-311-034-20	Sequence 20, Appl
20	3580.5	8.7	26926	9	US-09-759-508B-2	Sequence 2, Appl
21	2629	6.4	4162	15	US-10-023-634-92	Sequence 92, Appl
22	2489	6.0	6642	15	US-10-369-493-5013	Sequence 5013, Ap
23	2447.5	5.9	548	14	US-10-307-019-8	Sequence 8, Appl
24	2415.5	5.9	548	14	US-10-307-019-9	Sequence 9, Appl
25	2324	5.6	3208	15	US-10-210-130-38	Sequence 38, Appl
26	2319	5.6	3268	15	US-10-379-381-2	Sequence 2, Appl
27	2270	5.5	3252	15	US-10-210-130-36	Sequence 36, Appl
28	2203	5.3	3262	15	US-10-379-381-4	Sequence 4, Appl
29	2135	5.2	2380	15	US-10-333-314-18	Sequence 18, Appl
30	2102	5.1	5635	15	US-10-451-168-78	Sequence 78, Appl
31	2085.5	5.1	3186	15	US-10-210-130-34	Sequence 34, Appl
32	2085	5.1	416	15	US-10-108-260A-4389	Sequence 4389, Ap
33	2072.5	5.0	5636	14	US-10-032-189-128	Sequence 128, App
34	2072.5	5.0	5636	15	US-10-120-801-72	Sequence 72, Appl
35	2072.5	5.0	5636	15	US-10-023-634-93	Sequence 93, Appl
36	2072.5	5.0	5636	16	US-10-408-765A-1895	Sequence 1895, Ap
37	2067	5.0	2231	15	US-10-379-381-5	Sequence 5, Appl
38	2067	5.0	2242	16	US-10-408-765A-793	Sequence 793, App
39	2048	5.0	5198	15	US-10-120-801-75	Sequence 75, Appl
40	2048	5.0	5198	15	US-10-369-493-6858	Sequence 6858, Ap
41	2048	5.0	5198	15	US-10-369-493-6860	Sequence 6860, Ap
42	2048	5.0	5198	15	US-10-346-863-34	Sequence 34, Appl
43	2045	5.0	390	15	US-10-425-114-37530	Sequence 37530, A
44	2043.5	5.0	5175	15	US-10-120-801-74	Sequence 74, Appl
45	2043.5	5.0	5175	15	US-10-369-493-6859	Sequence 6859, Ap

ALIGNMENTS

RESULT 1
US-10-077-130-5
; Sequence 5, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; TITLE OF INVENTION: Members and Uses Therefor
; FILE REFERENCE: MPI2001-047PIRCP1(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; PRIOR FILING DATE: 2001-02-15
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 7968
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-077-130-5

Query Match 100.0%; Score 41273; DB 13; Length 7968;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7968; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDQPFSGARFLTRPKAFVSVGKDATLSQIVGNPTPQVSWKDDQPVVTAGARFLAQ 60
Db 1 MDQPFSGARFLTRPKAFVSVGKDATLSQIVGNPTPQVSWKDDQPVVTAGARFLAQ 60
Qy 61 DGDYRLTILDALGDSQYVCRNNAI GEAFVAVGLQVDAEAAACAOAPHPLLRPTSR 120
Db 61 DGDYRLTILDALGDSQYVCRNNAI GEAFVAVGLQVDAEAAACAOAPHPLLRPTSR 120
Qy 121 VREGSEATFCRCVGSPPAVSWKDGRLGEGDGPVRVRELGEASALRIEAAAPROG 180

Db 121 VREGSEATFRRCRVGSPRPAPVWSKQDGRRLRCEPDGPRVVRVEELGEASALRIRAAARPRDGG 180
Qy 181 TYEVRAENPLGASAAAALVDSADATSRPGTSTAALLAHILORREANRAGAPASPP 240
Db 181 TYEVRAENPLGASAAAALVDSADATSRPGTSTAALLAHILORREANRAGAPASPP 240
Qy 241 STGTRTCTVTGEGKHARLSCVYTGEPKPEYVWKDQGLVTEGRRHVVEDAQENFVLKILF 300
Db 241 STGTRTCTVTGEGKHARLSCVYTGEPKPEYVWKDQGLVTEGRRHVVEDAQENFVLKILF 300
Qy 301 CKSDRGLYTCTASNLVQGYSSVLVVRREPAPVFKRLQDLEVRKESATFLCEVPQPS 360
Db 301 CKSDRGLYTCTASNLVQGYSSVLVVRREPAPVFKRLQDLEVRKESATFLCEVPQPS 360
Qy 361 TEAAWFKEETRLWASAKYGIEEGTERRLTVRNVSAADDDAVYCETPEGSRTVAELAVOG 420
Db 361 TEAAWFKEETRLWASAKYGIEEGTERRLTVRNVSAADDDAVYCETPEGSRTVAELAVOG 420
Qy 421 NLLRKLPRKTAVRVGDGTAMFCVELAVPVGPHVWLRNQEEVWAGGRVAISAEGTRHRTLIS 480
Db 421 NLLRKLPRKTAVRVGDGTAMFCVELAVPVGPHVWLRNQEEVWAGGRVAISAEGTRHRTLIS 480
Qy 481 QCLEDVGQVAFWAGDCOTSTRFCVSAPRKPLOPPVDVVKARMESSVILSNSPPHGE 540
Db 481 QCLEDVGQVAFWAGDCOTSTRFCVSAPRKPLOPPVDVVKARMESSVILSNSPPHGE 540
Qy 541 RPVTIDGYLVEKKLGTYYTIRCHEAEWVATPELTIVADVAEENGFQFRVSALNSFGQSPY 600
Db 541 RPVTIDGYLVEKKLGTYYTIRCHEAEWVATPELTIVADVAEENGFQFRVSALNSFGQSPY 600
Qy 601 LEPPGTVHLAPKLVATPLKAVQAVEGGEVTSVDLTVASAGWFJLDGQALKASSVYEIH 660
Db 601 LEPPGTVHLAPKLVATPLKAVQAVEGGEVTSVDLTVASAGWFJLDGQALKASSVYEIH 660
Qy 661 CDRTRHTLTIREVPASLHGAQLKFVANGIESSIRMEVRAAPGLTANKPAAAAREVLARL 720
Db 661 CDRTRHTLTIREVPASLHGAQLKFVANGIESSIRMEVRAAPGLTANKPAAAAREVLARL 720
Qy 721 HEEAQLLAEISDQAAAVTWLKDGRTLSPGPKYEVQASARRVLLVRDVARDDAGLYECVS 780
Db 721 HEEAQLLAEISDQAAAVTWLKDGRTLSPGPKYEVQASARRVLLVRDVARDDAGLYECVS 780
Qy 781 RGRITAYQLSVQGLARFLHKNWAGSCVDVAGGPAQFECETSEAHVHVHWYKDGHELGHS 840
Db 781 RGRITAYQLSVQGLARFLHKNWAGSCVDVAGGPAQFECETSEAHVHVHWYKDGHELGHS 840
Qy 841 GERFLOEDVGTTRHLVAATVTTRQDEGTYSCRVEDSDVDFRLRVSEPKVVFVFAKEQLARRKL 900
Db 841 GERFLOEDVGTTRHLVAATVTTRQDEGTYSCRVEDSDVDFRLRVSEPKVVFVFAKEQLARRKL 900
Qy 901 QAEAGASATLSCEVAQAQTEVTWYKDGKLSLSSSKVCMEATGCTTRRLVVQAGQADAGEY 960
Db 901 QAEAGASATLSCEVAQAQTEVTWYKDGKLSLSSSKVCMEATGCTTRRLVVQAGQADAGEY 960
Qy 961 SCEAGQRLSFHLDVYKPKVFAKDQVAHSEVQAEAGANATLSCEVAQAQEVWVWYKDGK 1020
Db 961 SCEAGQRLSFHLDVYKPKVFAKDQVAHSEVQAEAGANATLSCEVAQAQEVWVWYKDGK 1020
Qy 1021 KLSSSLKVHVEAKGRRRLVWQAGTKTDAGDYSCBARGQVRSFRLHITPEPOMFAKEQSV 1080
Db 1021 KLSSSLKVHVEAKGRRRLVWQAGTKTDAGDYSCBARGQVRSFRLHITPEPOMFAKEQSV 1080
Qy 1081 HNEVQAEAGASAMLSCEVAQAQTEVTWYKDGKLSLSSSKVCMEVKGCTTRRLVLVQAGKAD 1140
Db 1081 HNEVQAEAGASAMLSCEVAQAQTEVTWYKDGKLSLSSSKVCMEVKGCTTRRLVLVQAGKAD 1140
Qy 1141 AGEYSCEAGGQVRSFHLHITBPKVFAKEQSVHNEVQAEAGTAMLSCEVAQAQTEVTWY 1200
Db 1141 AGEYSCEAGGQVRSFHLHITBPKVFAKEQSVHNEVQAEAGTAMLSCEVAQAQTEVTWY 1200
Qy 1201 KDGKLSLSSSKVMEVKGCTTRRLVVQGVKADAGEYSCBARGQVRSFOLHITPEKAVFAK 1260
Db 1201 KDGKLSLSSSKVMEVKGCTTRRLVVQGVKADAGEYSCBARGQVRSFOLHITPEKAVFAK 1260

Qy 1261 EOLVNEVNRTEAGASATLSCEVAQAQTEVTWYKDGKLSLSSSKVRIEAGCMRQLVVQQA 1320
Db 1261 EOLVNEVNRTEAGASATLSCEVAQAQTEVTWYKDGKLSLSSSKVRIEAGCMRQLVVQQA 1320
Qy 1321 GQADAGEYTCBAGGQRLSFHLDVSEPKAVFAKEQLAHRKVQAEAGAIATLSCEVAQAQTE 1380
Db 1321 GQADAGEYTCBAGGQRLSFHLDVSEPKAVFAKEQLAHRKVQAEAGAIATLSCEVAQAQTE 1380
Qy 1381 VTWYKDGKLSLSSSKVMEAVGCTTRRLVVQACQADTGEYSCBAGGQRLSFSLDVABPKV 1440
Db 1381 VTWYKDGKLSLSSSKVMEAVGCTTRRLVVQACQADTGEYSCBAGGQRLSFSLDVABPKV 1440
Qy 1441 VPAKEOPVHREVQAGAGASTTILSCEVAQAQTEVTWYKDGKLSLSSSKVMEAVGCTTRRLV 1500
Db 1441 VPAKEOPVHREVQAGAGASTTILSCEVAQAQTEVTWYKDGKLSLSSSKVMEAVGCTTRRLV 1500
Qy 1501 VQAGQADAGEYSCBAGGQRLSFHLDVSEPKAVFAKEQBPASREVQAEAGTATLSCEVAQ 1560
Db 1501 VQAGQADAGEYSCBAGGQRLSFHLDVSEPKAVFAKEQBPASREVQAEAGTATLSCEVAQ 1560
Qy 1561 AQTEVTWYKDGKLSLSSSKVMEAVGCTTRRLVVQAEQADAGEYSCBAGQRLSFHLHVA 1620
Db 1561 AQTEVTWYKDGKLSLSSSKVMEAVGCTTRRLVVQAEQADAGEYSCBAGQRLSFHLHVA 1620
Qy 1621 EPKVVFPAKEQPAHREVQAEAGASATLSCEVAQAQTEVTWYKDGKLSLSSSKVREAVGCT 1680
Db 1621 EPKVVFPAKEQPAHREVQAEAGASATLSCEVAQAQTEVTWYKDGKLSLSSSKVREAVGCT 1680
Qy 1681 RRLVVQAGQADAGEYSCBAGGQRLSFRLHVAEILEPOISERPCCRREPLVVKHEHEDILTA 1740
Db 1681 RRLVVQAGQADAGEYSCBAGGQRLSFRLHVAEILEPOISERPCCRREPLVVKHEHEDILTA 1740
Qy 1741 TLATPSAATVTWLKDGVEIRRSKXHEHTASQGDTHTLTVHGAQVLDLSAIYSCRVAEGQDF 1800
Db 1741 TLATPSAATVTWLKDGVEIRRSKXHEHTASQGDTHTLTVHGAQVLDLSAIYSCRVAEGQDF 1800
Qy 1801 PVQVEEVAAXFCRLLEPVCGLGCTVTLACELSPACAEVVMRCGNTOPRVKGFQMVVAG 1860
Db 1801 PVQVEEVAAXFCRLLEPVCGLGCTVTLACELSPACAEVVMRCGNTOPRVKGFQMVVAG 1860
Qy 1861 PVRSRLTVGLRAEDAGEYVCESRDDHTSAQLTVSVPRVVKFMSGLSTVVAEEGGEATFQC 1920
Db 1861 PVRSRLTVGLRAEDAGEYVCESRDDHTSAQLTVSVPRVVKFMSGLSTVVAEEGGEATFQC 1920
Qy 1921 VVSPSDVAVVWFRDGLALOPSEKFAISQSGASHLTSIDLVLLEDAGQITVVEAGAGSSAA 1980
Db 1921 VVSPSDVAVVWFRDGLALOPSEKFAISQSGASHLTSIDLVLLEDAGQITVVEAGAGSSAA 1980
Qy 1981 LRVREAPVLFKKKLEPQTVTEERSSVTLEVELTRPWPPELWTRNATALAPGKNVEIHAEGA 2040
Db 1981 LRVREAPVLFKKKLEPQTVTEERSSVTLEVELTRPWPPELWTRNATALAPGKNVEIHAEGA 2040
Qy 2041 RHRLVLHNVGFADRGFPFGCETPDDKTQAKLTIVEMROVRLVIRGLOQVAEAREQGTATMEVOL 2100
Db 2041 RHRLVLHNVGFADRGFPFGCETPDDKTQAKLTIVEMROVRLVIRGLOQVAEAREQGTATMEVOL 2100
Qy 2101 SHADVGSWTRDGLRFQCGPTCHILAVRGPMTHTLTLGLRPEDSGLMVFKAEGVHTSARLV 2160
Db 2101 SHADVGSWTRDGLRFQCGPTCHILAVRGPMTHTLTLGLRPEDSGLMVFKAEGVHTSARLV 2160
Qy 2161 VTPELVSFRPLQDVVVTTEKEKVTLECELSRPNVDVRLKDGVELRAGKTMIAAQGACR 2220
Db 2161 VTPELVSFRPLQDVVVTTEKEKVTLECELSRPNVDVRLKDGVELRAGKTMIAAQGACR 2220
Qy 2221 SLTIYRCEPADQGVVCDADHAQSSASVKVQCGRTTLYIRRVRLAEDAGBIQVFAENAESR 2280
Db 2221 SLTIYRCEPADQGVVCDADHAQSSASVKVQCGRTTLYIRRVRLAEDAGBIQVFAENAESR 2280
Qy 2281 AQLRVKELPVTILVRPLRDKIAMKEKIRGVLECCQVSRASQVRWFKGSOELQPGKYELVSD 2340
Db 2281 AQLRVKELPVTILVRPLRDKIAMKEKIRGVLECCQVSRASQVRWFKGSOELQPGKYELVSD 2340

QY	2341	GLYRKLIIISDVHAEDEDTYTCAGDVKTSQAQFVBEQSIITIVRGLQDVTVMPEPAPWREC	2400
Db	2341	GLYRKLIIISDVHAEDEDTYTCAGDVKTSQAQFVBEQSIITIVRGLQDVTVMPEPAPWREC	2400
QY	2401	ETSIPIVPPKWLKTKTVLQAGNVGLBOEGTVHRLMLRRTCTSTMTGPVHFTVVGKSRSSA	2460
Db	2401	ETSIPIVPPKWLKTKTVLQAGNVGLBOEGTVHRLMLRRTCTSTMTGPVHFTVVGKSRSSA	2460
QY	2461	RLVVSDDIPVLTPLPEPTKGBRELQSVLS CDP RPAPKAVQMYKDDTPLSPSKFKPMSLEG	2520
Db	2461	RLVVSDDIPVLTPLPEPTKGBRELQSVLS CDP RPAPKAVQMYKDDTPLSPSKFKPMSLEG	2520
QY	2521	QMAELRIILRLMPADAGVVRQAGSAHSSTEVTVAREVTVTGPLQDAEATEGWSA FSC E	2580
Db	2521	QMAELRIILRLMPADAGVVRQAGSAHSSTEVTVAREVTVTGPLQDAEATEGWSA FSC E	2580
QY	2581	LSHEDEEVEWSLNGMPLNDSPHEISHKGRHRTLVLKSIQRADAGIVRASSILKVSTSA RL	2640
Db	2581	LSHEDEEVEWSLNGMPLNDSPHEISHKGRHRTLVLKSIQRADAGIVRASSILKVSTSA RL	2640
QY	2641	EVRVPVVFVKALDDLSAERGTALQCEVSDPEAHVVRKDGVOQLGPSDKYDFLHTAGT	2700
Db	2641	EVRVPVVFVKALDDLSAERGTALQCEVSDPEAHVVRKDGVOQLGPSDKYDFLHTAGT	2700
QY	2701	RGLVVHVSYPDAGLYTCHVGSSEETRARVRVHDLHVGITKRLKTMVLEGESCSPECVLS	2760
Db	2701	RGLVVHVSYPDAGLYTCHVGSSEETRARVRVHDLHVGITKRLKTMVLEGESCSPECVLS	2760
QY	2761	HESASDPAMVTGGKTVGSSSRFOATROQRKYILVVREAAPSDAGEVVFVSRGLTSKASL	2820
Db	2761	HESASDPAMVTGGKTVGSSSRFOATROQRKYILVVREAAPSDAGEVVFVSRGLTSKASL	2820
QY	2821	IVRERPAAI IKPLEDQWVAPGSDVBLRCELRSAGTPVHVLKDRKAI RKSQKYDVVCEGTM	2880
Db	2821	IVRERPAAI IKPLEDQWVAPGSDVBLRCELRSAGTPVHVLKDRKAI RKSQKYDVVCEGTM	2880
QY	2881	AMLVIRGASLKDAGEYTCVEVASKSTASHVEEKANCFTTELNTLOVEKGTAVFTCKTE	2940
Db	2881	AMLVIRGASLKDAGEYTCVEVASKSTASHVEEKANCFTTELNTLOVEKGTAVFTCKTE	2940
QY	2941	HPAATVTVWRKGLLELRASGKQHPQSEGLTLRLTISALEKADSDTYTCIDIGQAQSAQLLV	3000
Db	2941	HPAATVTVWRKGLLELRASGKQHPQSEGLTLRLTISALEKADSDTYTCIDIGQAQSAQLLV	3000
QY	3001	QGRRVHIIIEDLVVDVQBGSSATFCRISPANYEPVHMFPLDKTPHLANELNEIDAQPGGY	3060
Db	3001	QGRRVHIIIEDLVVDVQBGSSATFCRISPANYEPVHMFPLDKTPHLANELNEIDAQPGGY	3060
QY	3061	HVLTIRQLALKDSGTIYFEAGDORASALRVTEKPSVFSRELTDATITTEGEDTLVVCETS	3120
Db	3061	HVLTIRQLALKDSGTIYFEAGDORASALRVTEKPSVFSRELTDATITTEGEDTLVVCETS	3120
QY	3121	TCDIPMCWTKDGKTLRGSAQCQSHQHEHRAQLLIITGATLQDSGRYKCEAGGACSSSIVRV	3180
Db	3121	TCDIPMCWTKDGKTLRGSAQCQSHQHEHRAQLLIITGATLQDSGRYKCEAGGACSSSIVRV	3180
QY	3181	HARVPRFOEALKOLEVLBEGAATLRCVLS SVAAPVKWCYGNVLRPDKYSLRQEGAMLE	3240
Db	3181	HARVPRFOEALKOLEVLBEGAATLRCVLS SVAAPVKWCYGNVLRPDKYSLRQEGAMLE	3240
QY	3241	LWVNRRLRQDSGRYSCSGEDQTTSATLTVTALPAQFVIGKLRNKEATEGATLRLCELSKT	3300
Db	3241	LWVNRRLRQDSGRYSCSGEDQTTSATLTVTALPAQFVIGKLRNKEATEGATLRLCELSKT	3300
QY	3301	APVEWRKGSSETLRDGDRLRQDQGANCELOIRGLAMVDAAEYSCVCGEERTSASITIRPM	3360
Db	3301	APVEWRKGSSETLRDGDRLRQDQGANCELOIRGLAMVDAAEYSCVCGEERTSASITIRPM	3360
QY	3361	PAHFIGRLRHQESIEGATATLRLCELSKAAPVEWRKRESLRDGDHRS LRQDGA VCELQIC	3420
Db	3361	PAHFIGRLRHQESIEGATATLRLCELSKAAPVEWRKRESLRDGDHRS LRQDGA VCELQIC	3420
QY	3421	GLAVADAGEYSCVCGEERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLWCBSKVAPVE	3480

Db	3421	GLAVADAGEYSCVCGEERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLWCBSKVAPVE	3480
QY	3481	WRKGPENLRDGDRLRQEGTRCELCQICGLAMADAGEYLCVCGQBERTSATLTIRALPARF	3540
Db	3481	WRKGPENLRDGDRLRQEGTRCELCQICGLAMADAGEYLCVCGQBERTSATLTIRALPARF	3540
QY	3541	IEDVKNQAREGATAVLOCELNSAAPVEWRKGSSETLRDGDRLRQDGT KCBLQIRGLAM	3600
Db	3541	IEDVKNQAREGATAVLOCELNSAAPVEWRKGSSETLRDGDRLRQDGT KCBLQIRGLAM	3600
QY	3601	ADTGEYSCVCGQBERTSAMLTVRALPIKFTTEGLRNEEATEGATAVLRCELSKAAPVEWRKGS	3660
Db	3601	ADTGEYSCVCGQBERTSAMLTVRALPIKFTTEGLRNEEATEGATAVLRCELSKAAPVEWRKGS	3660
QY	3661	HETLRDGDRLRQDQGANCELOIRGLVAEDAGEYLCMCGKERTSAMLTVRAMPSKFIEGL	3720
Db	3661	HETLRDGDRLRQDQGANCELOIRGLVAEDAGEYLCMCGKERTSAMLTVRAMPSKFIEGL	3720
QY	3721	RNEEATEGDTATLWCELSKAAPVEWRKCHETLRDGDHRS LRQDGRSCELOIRGLAVVDAG	3780
Db	3721	RNEEATEGDTATLWCELSKAAPVEWRKCHETLRDGDHRS LRQDGRSCELOIRGLAVVDAG	3780
QY	3781	EYSCVCGQBERTSATLTVRALPARFIEDVKNQAREGATAVLOCELSKAAPVEWRKGSSETL	3840
Db	3781	EYSCVCGQBERTSATLTVRALPARFIEDVKNQAREGATAVLOCELSKAAPVEWRKGSSETL	3840
QY	3841	RGGRYSLRQDGTCELCQIHGLSVADTGEYSCVCGQBERTSATLTVRAPQVPVPREPLQSLQ	3900
Db	3841	RGGRYSLRQDGTCELCQIHGLSVADTGEYSCVCGQBERTSATLTVRAPQVPVPREPLQSLQ	3900
QY	3901	ABEGSTATLQCELSBPTATVWVSKGLOIQANGREPRELQGTAEVLQDLOREDTGEYT	3960
Db	3901	ABEGSTATLQCELSBPTATVWVSKGLOIQANGREPRELQGTAEVLQDLOREDTGEYT	3960
QY	3961	CTCGSQATSATLTVTAAAPVRFRLRLOHQBEGGTAHLCCELSRAGASVENWKGSLQLPF	4020
Db	3961	CTCGSQATSATLTVTAAAPVRFRLRLOHQBEGGTAHLCCELSRAGASVENWKGSLQLPF	4020
QY	4021	CAYQMVQDQGAALLLVRGVEQEDAGDYTCDTGHTQSMASLSVRVPKFKTRLOSLEQE	4080
Db	4021	CAYQMVQDQGAALLLVRGVEQEDAGDYTCDTGHTQSMASLSVRVPKFKTRLOSLEQE	4080
QY	4081	TGDIARLCCQLSDABESGAVOWMLKEGVELHAGPKYEMRSQGNATRELLIHQLEBAKTGEYA	4140
Db	4081	TGDIARLCCQLSDABESGAVOWMLKEGVELHAGPKYEMRSQGNATRELLIHQLEBAKTGEYA	4140
QY	4141	CVTGQKTAASLRVTPEPTIVRGLVDAEVTADDEVEFSCEVSRAGATGVQWCLQGLPIQ	4200
Db	4141	CVTGQKTAASLRVTPEPTIVRGLVDAEVTADDEVEFSCEVSRAGATGVQWCLQGLPIQ	4200
QY	4201	SNEVTEVAVRDGRHITLRLKGVTPEDAGTVSFHGLGNHASSAQLTVRAPEVTILEPLQDVQ	4260
Db	4201	SNEVTEVAVRDGRHITLRLKGVTPEDAGTVSFHGLGNHASSAQLTVRAPEVTILEPLQDVQ	4260
QY	4261	LSSEGODASFCQRLSRASQBARWALGGVPLQANENMDITVREGTTLHLTLHKVTLEDACT	4320
Db	4261	LSSEGODASFCQRLSRASQBARWALGGVPLQANENMDITVREGTTLHLTLHKVTLEDACT	4320
QY	4321	VSEFHVGTCCSSAQLKVTAKNVTVRGLENVLEALEGGEALFECCOLSOPEVAHAHTWLLDDBPV	4380
Db	4321	VSEFHVGTCCSSAQLKVTAKNVTVRGLENVLEALEGGEALFECCOLSOPEVAHAHTWLLDDBPV	4380
QY	4381	RTSENAEYVVPFENGRLHLLLNLRKPQDSRCRTVFLAGDMVTSAFITVRGWRLEILEPLKN	4440
Db	4381	RTSENAEYVVPFENGRLHLLLNLRKPQDSRCRTVFLAGDMVTSAFITVRGWRLEILEPLKN	4440
QY	4441	AAVRAGAQAARFTCTLSEAVPVGEASWYINGAAVQDDSDTWTADGSHOALLLRSAQPHH	4500
Db	4441	AAVRAGAQAARFTCTLSEAVPVGEASWYINGAAVQDDSDTWTADGSHOALLLRSAQPHH	4500
QY	4501	AGEVTFACEDAVASARLTVLGLPDPPEAEVVAHSSHTVTLSSWAPMSDGGGLGCRVYE	4560

Db 4501 AGEVTFACRDVAASARLTVLGLPDPPEDAEVVAHSSHTVTL5WAAPMSDGGGLCGYRVE 4560
QY VKEGATGQWRLCHELVPGPECVVDGLAPGETYFRVAAGVPVGAGEPVHLPTQVRLAEP 4620
Db VKEGATGQWRLCHELVPGPECVVDGLAPGETYFRVAAGVPVGAGEPVHLPTQVRLAEP 4620
QY KPVPPOPSAPESRQVAAGSDVLSLELVVAEAGEVIMHKWMERIQPGREFVVSQGRQQML 4680
Db KPVPPOPSAPESRQVAAGSDVLSLELVVAEAGEVIMHKWMERIQPGREFVVSQGRQQML 4680
QY VIKGFTAEODGEGYHCGLAGOSTCPAAATFOVALSPASVDEAPQPSLPPEAAQEGDLHLW 4740
Db VIKGFTAEODGEGYHCGLAGOSTCPAAATFOVALSPASVDEAPQPSLPPEAAQEGDLHLW 4740
QY EALARKRMSRPTLDSISELPEEDGRSQRLLPOEAEEVAPDLSEGYSTADELARTGDADL 4800
Db EALARKRMSRPTLDSISELPEEDGRSQRLLPOEAEEVAPDLSEGYSTADELARTGDADL 4800
QY SHTSSDDESRACTPSLVTVYLKKAGRPGTSPLASKVGAPAAPSVKPOQOQEPPLAARVPLG 4860
Db SHTSSDDESRACTPSLVTVYLKKAGRPGTSPLASKVGAPAAPSVKPOQOQEPPLAARVPLG 4860
QY DLSTKDLGDPSPMDKAQKIAQAFKGYKVRKEMKQOEGPMFSTFGDTAEQVGDALRLBCV 4920
Db DLSTKDLGDPSPMDKAQKIAQAFKGYKVRKEMKQOEGPMFSTFGDTAEQVGDALRLBCV 4920
QY VASKADVRARWLKDGVELTDGRHHIDQLGDGTC5LLIAGLDRADAGCVTCOVSNKFGQV 4980
Db VASKADVRARWLKDGVELTDGRHHIDQLGDGTC5LLIAGLDRADAGCVTCOVSNKFGQV 4980
QY THSACVWVSGSSEAE5SGGBELDAFRAARLHRLFTKSPAEVSD5ELFLSADEGPA 5040
Db THSACVWVSGSSEAE5SGGBELDAFRAARLHRLFTKSPAEVSD5ELFLSADEGPA 5040
QY EPEEPADWQTYREDEHFICIRFEALTEARQAVTRQEMPATLIGIVEIKLVEQGRPRVEM 5100
Db EPEEPADWQTYREDEHFICIRFEALTEARQAVTRQEMPATLIGIVEIKLVEQGRPRVEM 5100
QY CISKETPAWPVPEPLPSLLTSDAAPVFLTELQNEVDGYPVSPDCVVVTGQPMPSVRWF 5160
Db CISKETPAWPVPEPLPSLLTSDAAPVFLTELQNEVDGYPVSPDCVVVTGQPMPSVRWF 5160
QY KDGLLEEDDHYMINEDQGGHQLIITAVVPADMGYTRCLAENSMGV5STKAEFLVDLTS 5220
Db KDGLLEEDDHYMINEDQGGHQLIITAVVPADMGYTRCLAENSMGV5STKAEFLVDLTS 5220
QY TDYDTAADATES5SYFSAQGYLSSREOGETSTTDEGQLPQVVEBLRDLQVAPGTRLAKF 5280
Db TDYDTAADATES5SYFSAQGYLSSREOGETSTTDEGQLPQVVEBLRDLQVAPGTRLAKF 5280
QY QLKVKGYAPRLYWFKDGOPLTASAHIRMTGKKILHTLEIISVTREDSGQVAAVYSNANG 5340
Db QLKVKGYAPRLYWFKDGOPLTASAHIRMTGKKILHTLEIISVTREDSGQVAAVYSNANG 5340
QY AAY5SARLLVRGDPDEEPKPAD5VHEQLVPPRMLERFTPKVKKGSSITFSVKVGEPRVP 5400
Db AAY5SARLLVRGDPDEEPKPAD5VHEQLVPPRMLERFTPKVKKGSSITFSVKVGEPRVP 5400
QY TVHMLREBAERGVLATGPDTPGYTVASSAQHSLVLLDVGRHQGTYYTCIASNAAGQALC 5460
Db TVHMLREBAERGVLATGPDTPGYTVASSAQHSLVLLDVGRHQGTYYTCIASNAAGQALC 5460
QY SASLHV5GLPKV5EQEKVKEALISTFLOGTTQAI5AQGLETFASPADLGGQKEEPLAAKE 5520
Db SASLHV5GLPKV5EQEKVKEALISTFLOGTTQAI5AQGLETFASPADLGGQKEEPLAAKE 5520
QY ALGHLSLAEVGTPEFLQKLT5QITEMVSAKITQAKLQVPGGSD5D5SKTP5ASPRHGRSR 5580
Db ALGHLSLAEVGTPEFLQKLT5QITEMVSAKITQAKLQVPGGSD5D5SKTP5ASPRHGRSR 5580
QY PSS5IOESS5E55DGDARGEIPDIYVVVTADYPLICAEQDAILR5GOV5EVLDAAHPLRW 5640
Db PSS5IOESS5E55DGDARGEIPDIYVVVTADYPLICAEQDAILR5GOV5EVLDAAHPLRW 5640

QY LVRTKPTKSSPSRQGW5SPAYLDRRLKLSPEWGAEEAPEFFPCEAVSEDEYKARLSSVIOE 5700
Db LVRTKPTKSSPSRQGW5SPAYLDRRLKLSPEWGAEEAPEFFPCEAVSEDEYKARLSSVIOE 5700
QY LLSSEQA5VBEELQFLQSHHLQHLERC5PHVPIAVAGOKAVIPRNVRDIDGRFHSS5FLOELQO 5760
Db LLSSEQA5VBEELQFLQSHHLQHLERC5PHVPIAVAGOKAVIPRNVRDIDGRFHSS5FLOELQO 5760
QY CDTDDDVAMCFTKNOAAFEQYLEFLVGRVOAESVVV5TAIOEFYK5YAE5EALLAGDPSOP 5820
Db CDTDDDVAMCFTKNOAAFEQYLEFLVGRVOAESVVV5TAIOEFYK5YAE5EALLAGDPSOP 5820
QY PPPPLQHYLEOP5ERVQRYQALLKELIRNKARNRNCALL5QAVAVVSALPQAE5NKLHV 5880
Db PPPPLQHYLEOP5ERVQRYQALLKELIRNKARNRNCALL5QAVAVVSALPQAE5NKLHV 5880
QY SLMENYPGTL5ALGEP5IRQGHFI5VWEGAPGARM5PK5GHN5RVFL5FRNHLV5ICKPR5D5RT 5940
Db SLMENYPGTL5ALGEP5IRQGHFI5VWEGAPGARM5PK5GHN5RVFL5FRNHLV5ICKPR5D5RT 5940
QY DTVS5V5PENMMK5L5S5IDLNDQ5VEGDD5RA5FVWQ5REDS5VRK5YLQ5ARTAI5IKSS5W5KEIC 6000
Db DTVS5V5PENMMK5L5S5IDLNDQ5VEGDD5RA5FVWQ5REDS5VRK5YLQ5ARTAI5IKSS5W5KEIC 6000
QY GIQORLAL5PVWR5PDP5FEE5LADCTAEL5GET5VK5LAC5RV5T5G5TP5K5VISW5YK5DG5K5AVQ5D5PHH 6060
Db GIQORLAL5PVWR5PDP5FEE5LADCTAEL5GET5VK5LAC5RV5T5G5TP5K5VISW5YK5DG5K5AVQ5D5PHH 6060
QY ILI5BDP5G5CAL5IL5D5LT5G5V5D5G5Q5Y5M5CP5A5A5G5C5TL5G5K5IL5V5Q5V5PR5F5V5N5K5VR5AS5PV 6120
Db ILI5BDP5G5CAL5IL5D5LT5G5V5D5G5Q5Y5M5CP5A5A5G5C5TL5G5K5IL5V5Q5V5PR5F5V5N5K5VR5AS5PV 6120
QY E5G5DAQ5F5CT5T5LEG5AP5Q5IR5W5YK5D5G5ALL5T5CN5K5F5QT5L5EP5SG5LL5V5L5V5RA5S5K5ED5GL5Y 6180
Db E5G5DAQ5F5CT5T5LEG5AP5Q5IR5W5YK5D5G5ALL5T5CN5K5F5QT5L5EP5SG5LL5V5L5V5RA5S5K5ED5GL5Y 6180
QY E5CEL5V5N5L5G5S5A5E5L5RI5Q5P5ML5QA5O5B5Q5H5RE5QL5VA5VED5T5TL5ER5AD5Q5E5V5S5VL5K5RL5LG 6240
Db E5CEL5V5N5L5G5S5A5E5L5RI5Q5P5ML5QA5O5B5Q5H5RE5QL5VA5VED5T5TL5ER5AD5Q5E5V5S5VL5K5RL5LG 6240
QY PKA5P5SG5T5D5LT5G5P5C5PR5G5AP5AL5Q5E5T5G5S5P5V5T5G5T5E5AP5V5PR5V5P5O5P5L5L5H5E5G5E5Q5B5PE 6300
Db PKA5P5SG5T5D5LT5G5P5C5PR5G5AP5AL5Q5E5T5G5S5P5V5T5G5T5E5AP5V5PR5V5P5O5P5L5L5H5E5G5E5Q5B5PE 6300
QY A5I5A5R5AQ5E5W5T5P5IR5M5E5GA5W5P5G5AG5T5G5ELL5W5D5V5H5SV5RE5T5T5Q5T5Y5Q5A5ID5T5H5T5AR5P5SMQ 6360
Db A5I5A5R5AQ5E5W5T5P5IR5M5E5GA5W5P5G5AG5T5G5ELL5W5D5V5H5SV5RE5T5T5Q5T5Y5Q5A5ID5T5H5T5AR5P5SMQ 6360
QY VT5IED5V5QA5QT5G5TA5Q5F5E5A5I5EG5DP5Q5P5V5T5W5YK5D5V5L5D5S5T5RL5S5Q5Q5EG5T5Y5S5L5V5L5R5H5VA 6420
Db VT5IED5V5QA5QT5G5TA5Q5F5E5A5I5EG5DP5Q5P5V5T5W5YK5D5V5L5D5S5T5RL5S5Q5Q5EG5T5Y5S5L5V5L5R5H5VA 6420
QY SKD5AG5V5T5CLA5Q5T5G5G5V5L5C5K5A5E5LL5V5L5G5D5N5E5P5S5E5K5Q5H5R5K5L5H5S5F5E5V5E5E5I5G5R5G5V5G 6480
Db SKD5AG5V5T5CLA5Q5T5G5G5V5L5C5K5A5E5LL5V5L5G5D5N5E5P5S5E5K5Q5H5R5K5L5H5S5F5E5V5E5E5I5G5R5G5V5G 6480
QY FV5K5E5V5Q5H5K5N5K5IL5C5A5K5F5I5P5L5R5S5T5RA5Q5A5V5E5R5D5I5A5L5S5H5P5L5V5T5G5L5D5Q5E5T5R5K5T5L5LI 6540
Db FV5K5E5V5Q5H5K5N5K5IL5C5A5K5F5I5P5L5R5S5T5RA5Q5A5V5E5R5D5I5A5L5S5H5P5L5V5T5G5L5D5Q5E5T5R5K5T5L5LI 6540
QY LEL5C5S5E5L5D5L5R5K5Y5K5V5T5E5A5E5V5K5VI5Q5I5Q5L5V5E5G5L5H5L5H5S5H5V5L5H5D5I5K5P5N5I5L5M5V5H5P5AR 6600
Db LEL5C5S5E5L5D5L5R5K5Y5K5V5T5E5A5E5V5K5VI5Q5I5Q5L5V5E5G5L5H5L5H5S5H5V5L5H5D5I5K5P5N5I5L5M5V5H5P5AR 6600
QY EDI5K5I5C5F5G5P5A5Q5N5IT5P5A5E5L5Q5F5S5Q5G5P5E5F5S5P5E5I5I5Q5N5P5V5E5A5D5I5W5A5G5V5I5S5L5T5C5S 6660
Db EDI5K5I5C5F5G5P5A5Q5N5IT5P5A5E5L5Q5F5S5Q5G5P5E5F5S5P5E5I5I5Q5N5P5V5E5A5D5I5W5A5G5V5I5S5L5T5C5S 6660
QY SP5F5A5G5E5D5R5A5T5L5N5V5L5G5R5V5S5W5S5P5M5A5H5L5S5E5A5K5D5F5I5K5A5T5L5Q5R5A5P5A5A5Q5C5L5S5H5P5W 6720
Db SP5F5A5G5E5D5R5A5T5L5N5V5L5G5R5V5S5W5S5P5M5A5H5L5S5E5A5K5D5F5I5K5A5T5L5Q5R5A5P5A5A5Q5C5L5S5H5P5W 6720

QY 6721 FLKSMPEAEHAFINTKOLKFLARSRWORSLSYKSLVMSRIPELLRGPPDSPSLGVAR 6780
Db 6721 FLKSMPEAEHAFINTKOLKFLARSRWORSLSYKSLVMSRIPELLRGPPDSPSLGVAR 6780
QY 6781 HLCRTDGGSSSSSSSSDNEAPAFARAKSLPPSPVTHSPHLLHPRGFLRPSASLPPEAEASE 6840
Db 6781 HLCRTDGGSSSSSSSSDNEAPAFARAKSLPPSPVTHSPHLLHPRGFLRPSASLPPEAEASE 6840
QY 6841 RSTEPAPASPEGAGPPAAQCCVPRHSVIRSLFTHQAGESPEHGALAPGSRHRPARRH 6900
Db 6841 RSTEPAPASPEGAGPPAAQCCVPRHSVIRSLFTHQAGESPEHGALAPGSRHRPARRH 6900
QY 6901 LLKGVIAGALPGLREPLMEHRLVEEAREEQATLLAKAPSFETALRLPASGTHLAPGH 6960
Db 6901 LLKGVIAGALPGLREPLMEHRLVEEAREEQATLLAKAPSFETALRLPASGTHLAPGH 6960
QY 6961 SHSLHDSFSTPRPSSSEACGEAQLRPSAPSGAPIRDMGHPQGSQKLPSTGCHPCTAQPE 7020
Db 6961 SHSLHDSFSTPRPSSSEACGEAQLRPSAPSGAPIRDMGHPQGSQKLPSTGCHPCTAQPE 7020
QY 7021 RPSDSPNGQAPFCHPQGSAPQSGSPHPAVAPCPGSPFPSPGCKEAPLVPSSPFLGQ 7080
Db 7021 RPSDSPNGQAPFCHPQGSAPQSGSPHPAVAPCPGSPFPSPGCKEAPLVPSSPFLGQ 7080
QY 7081 PQAPPAPAKASPPLDSKMGCGDISLPGRPKPCSSPGSASQASQSSVSLRVGSSQVCT 7140
Db 7081 PQAPPAPAKASPPLDSKMGCGDISLPGRPKPCSSPGSASQASQSSVSLRVGSSQVCT 7140
QY 7141 EPGSLDAEGMTQEAEDLSDSPTTLQRPQEQATMRKFSLGGRGGYAGVAGYGTFAFGDA 7200
Db 7141 EPGSLDAEGMTQEAEDLSDSPTTLQRPQEQATMRKFSLGGRGGYAGVAGYGTFAFGDA 7200
QY 7201 GGMLCQGPMMWARIANAVQSSSEEEQEARAESQSEEQEARAESPLQVARSAPVEVGRA 7260
Db 7201 GGMLCQGPMMWARIANAVQSSSEEEQEARAESQSEEQEARAESPLQVARSAPVEVGRA 7260
QY 7261 PTRSSPEPTPMDIGQVSLVQIRDSGDAEADTTSLDISEVDPAVNLSDLYDKYLPF 7320
Db 7261 PTRSSPEPTPMDIGQVSLVQIRDSGDAEADTTSLDISEVDPAVNLSDLYDKYLPF 7320
QY 7321 EFMI FRKVPKSAQPPPPMAEEELAEPEPTWMPGELGPHAGLEITEESDDVALLAE 7380
Db 7321 EFMI FRKVPKSAQPPPPMAEEELAEPEPTWMPGELGPHAGLEITEESDDVALLAE 7380
QY 7381 AAVGRKRWSSPRLPFPGRHPLDDEPAELGLRERVKASVEHISRLKGRPEGLEKEG 7440
Db 7381 AAVGRKRWSSPRLPFPGRHPLDDEPAELGLRERVKASVEHISRLKGRPEGLEKEG 7440
QY 7441 PRRKPGGLASFLSLGSKWDRAPTFLRLSDVTLGQSVTLACQVSAQPAQAATWSKDG 7500
Db 7441 PRRKPGGLASFLSLGSKWDRAPTFLRLSDVTLGQSVTLACQVSAQPAQAATWSKDG 7500
QY 7501 APLESSRVLISATLKNFOLLTLVVVAEDLGVTTCVSNALGVTVTGVLKKAERPSSS 7560
Db 7501 APLESSRVLISATLKNFOLLTLVVVAEDLGVTTCVSNALGVTVTGVLKKAERPSSS 7560
QY 7561 PCPDIGEVYADGVLLVWKPVSYPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRG 7620
Db 7561 PCPDIGEVYADGVLLVWKPVSYPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSRG 7620
QY 7621 TYTFTACVSKAGMGFPYSPSEQVLLGGPSHLASEESQGRSAQPLPSTKTFATQIQOR 7680
Db 7621 TYTFTACVSKAGMGFPYSPSEQVLLGGPSHLASEESQGRSAQPLPSTKTFATQIQOR 7680
QY 7681 GRFSVVRQWEKASGRALAALKIIPVHPKDKTAVLREYALKGLRPHLAQHLAAVLSPRH 7740
Db 7681 GRFSVVRQWEKASGRALAALKIIPVHPKDKTAVLREYALKGLRPHLAQHLAAVLSPRH 7740
QY 7741 LVLLLELCSGPPELLPCLAPRASYSSEVKDYLMQMSATQYLHNQHILHLDLRSNMIIIT 7800
Db 7741 LVLLLELCSGPPELLPCLAPRASYSSEVKDYLMQMSATQYLHNQHILHLDLRSNMIIIT 7800
QY 7801 EYNLLKVVLDLGNASQSLSEKVLPSDKFKDYLETMAPELLEGQCAVPQTDIWAIGVATIM 7860

RESULT 2

US-10-080-334-290

; Sequence 290, Application US/10080334

; Publication No. US20040002584A1

; GENERAL INFORMATION:

; APPLICANT: Pena, Carol E. A.

; APPLICANT: Shimkets, Richard A

; APPLICANT: Li, Li

; APPLICANT: Shenoy, Sureeh G

; APPLICANT: Kekuda, Rameesh

; APPLICANT: Spvtek, Kimberly A

; APPLICANT: Vernet, Corine A. M.

; APPLICANT: Malyankar, Uriel M

; APPLICANT: Guo, Xiaojia

; APPLICANT: Gusev, Vladimir Y

; APPLICANT: Casman, Stacie J

; APPLICANT: Boldog, Ferenc L

; APPLICANT: Furtak, Katarzyna

; APPLICANT: Tchernev, Velizar T

; APPLICANT: Patturajan, Meera

; APPLICANT: Gangolli, Esha A

; APPLICANT: Padigaru, Muralidhara

; APPLICANT: Liu, Xiaohong

; APPLICANT: Baumgartner, Jason C.

; APPLICANT: Gerlach, Valerie

; APPLICANT: Zethusen, Bryan D

; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of

; FILE REFERENCE: 21402-275

; CURRENT FILING DATE: 2002-02-21

; PRIOR APPLICATION NUMBER: US/10/080,334

; PRIOR FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: 60/270,523

; PRIOR FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: 60/322,712

; PRIOR FILING DATE: 2001-09-17

; PRIOR APPLICATION NUMBER: 60/311,980

; PRIOR FILING DATE: 2001-08-13

; PRIOR APPLICATION NUMBER: 60/330,307

; PRIOR FILING DATE: 2001-10-18

; PRIOR APPLICATION NUMBER: 60/278,796

; PRIOR FILING DATE: 2001-03-26

; PRIOR APPLICATION NUMBER: 60/281,521

; PRIOR FILING DATE: 2001-04-04

; PRIOR APPLICATION NUMBER: 60/276,677

; PRIOR FILING DATE: 2001-03-16

; PRIOR APPLICATION NUMBER: 60/311,595

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: 60/270,220

; PRIOR FILING DATE: 2001-02-21

; PRIOR APPLICATION NUMBER: 60/274,295

; PRIOR FILING DATE: 2001-03-08

; PRIOR APPLICATION NUMBER: 60/318,526

; PRIOR FILING DATE: 2001-09-10

; PRIOR APPLICATION NUMBER: 60/286,548

; PRIOR FILING DATE: 2001-04-25

; PRIOR APPLICATION NUMBER: 60/291,765

; PRIOR FILING DATE: 2001-05-17

; PRIOR APPLICATION NUMBER: 60/270,797

; PRIOR FILING DATE: 2001-02-23

; PRIOR APPLICATION NUMBER: 60/276,400

; PRIOR FILING DATE: 2001-03-16									
; PRIOR APPLICATION NUMBER: 60/270,810									
; PRIOR FILING DATE: 2001-02-23									
; NUMBER OF SEQ ID NOS: 388									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 290									
; LENGTH: 6620									
; TYPE: PRT									
; ORGANISM: Homo sapiens									
US-10-080-334-290									
Query Match 77.9%; Score 32134; DB 15; Length 6620;									
Best Local Similarity 98.1%; Pred. No. 0;									
Matches 6250; Conservative 15; Mismatches 57; Indels 48; Gaps 9;									
Qy	1	MDQPFSGAPRLTRPKAFVSVGKDATLSQOIVGNPTPOVSWEXDQOPVTAGARFLAQ	60						
Db	1	MDQPFSGAPRLTRPKAFVSVGKDATLSQOIVGNPTPOVSWEXDQOPVTAGARFLAQ	60						
Qy	61	DGDLRLTILDALGDSGOYVCRARNAIGEAFAAVGLQVDAAEACAEQAPHLRPTISR	120						
Db	61	DGDLRLTILDALGDSGOYVCRARNAIGEAFAAVGLQVDAAEACAEQAPHLRPTISR	120						
Qy	121	VREGSEATRCRVGSGPPAVSWKDGRRLLGPDGPRVRVBEELGEASALIRTRAARPRDGG	180						
Db	121	VREGSEATRCRVGSGPPAVSWKDGRRLLGPDGPRVRVBEELGEASALIRTRAARPRDGG	180						
Qy	181	TYEVRAENPLGAASAAALVDSDAADTASRPGTSTAALLAHLORRREARAEAPASPP	240						
Db	181	TYEVRAENPLGAASAAALVDSDAADTASRPGTSTAALLAHLORRREARAEAPASPP	240						
Qy	241	STGTRTCTVTEGKHARLSCYVTGEPKPTWKKDQQLVTGRRHVYVEDAQNFLVILF	300						
Db	241	STGTRTCTVTEGKHARLSCYVTGEPKPTWKKDQQLVTGRRHVYVEDAQNFLVILF	300						
Qy	301	CKQDRGLYTCASNLVGQTSVVLVVRREPAPVFKRLQDLVREKESATFLCEVPQPS	360						
Db	301	CKQDRGLYTCASNLVGQTSVVLVVRREPAPVFKRLQDLVREKESATFLCEVPQPS	360						
Qy	361	TEAAMFKETRLWASAKYIEEGTERLTVRNVSADDDAVYICETPEGSRTVAELAVQG	420						
Db	361	TEAAMFKETRLWASAKYIEEGTERLTVRNVSADDDAVYICETPEGSRTVAELAVQG	420						
Qy	421	NLLRLPKRTAVRVGDTAMFCVELAVPVGVHVLNQEVEVAGGRVAISAEGRHTLTIS	480						
Db	421	NLLRLPKRTAVRVGDTAMFCVELAVPVGVHVLNQEVEVAGGRVAISAEGRHTLTIS	480						
Qy	481	QCCLLEDVGVAFMAGDCQTSRFCVSAAPRKPPLQPPDPVVKARMESSVILSWSPPHGE	540						
Db	481	QCCLLEDVGVAFMAGDCQTSRFCVSAAPRKPPLQPPDPVVKARMESSVILSWSPPHGE	540						
Qy	541	RPVTIDGYLVEKKLGTWTWIRCHEAEWVATPELTVADVAEEGNFQPRVSALNSFGOSPY	600						
Db	541	RPVTIDGYLVEKKLGTWTWIRCHEAEWVATPELTVADVAEEGNFQPRVSALNSFGOSPY	600						
Qy	601	LEFPGTVHLAPKLAVRTPLKAVQAVEGGEVTFSDTLTVASAGFWFLDQALKASSVYIEH	660						
Db	601	LEFPGTVHLAPKLAVRTPLKAVQAVEGGEVTFSDTLTVASAGFWFLDQALKASSVYIEH	660						
Qy	661	CDTRHTLTIREVPASLHGAQIKFVANGIESSIRMEVRAAAGLTANKPAAAAAEVLARL	720						
Db	661	CDTRHTLTIREVPASLHGAQIKFVANGIESSIRMEVRAAAGLTANKPAAAAAEVLARL	720						
Qy	721	HEEAQLLAELSDQAAAVTWLKDGRTLSPGPKYEQVQASAGRRVLLVRDVARDDAGLYECVS	780						
Db	721	HEEAQLLAELSDQAAAVTWLKDGRTLSPGPKYEQVQASAGRRVLLVRDVARDDAGLYECVS	780						
Qy	781	RGRIAYQLSVQGLARFLHDKMAGSCVDAVAGGPAQPECETSEARVHVHMYKGMELGHS	840						
Db	781	RGRIAYQLSVQGLARFLHDKMAGSCVDAVAGGPAQPECETSEARVHVHMYKGMELGHS	840						
Qy	841	GERFLOEDVGTRHRLVAATVTRQDEGTYSCRVGSDVDFRLRVSEPKVVFVFAKEQLARRKL	900						

QY 1981 LRVRAPVLFKKLBPQTVBERSSVTLEVLTRPMPPELRWTRNATALAPGKNVEIHABGA 2040
DB 1981 LRVRAPVLFKKLBPQTVBERSSVTLEVLTRPMPPELRWTRNATALAPGKNVEIHABGA 2040
QY 2041 RHRLVLHNVGADRGFFGCEPDDKTQAKLTVMEMQVRLVRGLQAVEAREOQTATMEVOL 2100
DB 2041 RHRLVLHNVGADRGFFGCEPDDKTQAKLTVMEMQVRLVRGLQAVEAREOQTATMEVOL 2100
QY 2101 SHADVDGSWTRDGLRFQOQPTCHLAVRGPMTLTLISGLRPEDSGLMVFKAEGVHTSARLV 2160
DB 2101 SHADVDGSWTRDGLRFQOQPTCHLAVRGPMTLTLISGLRPEDSGLMVFKAEGVHTSARLV 2160
QY 2161 VTELPVFSRPLQDVVTTEKEKVTLECELSRPNVDVRLKOGVELRAGKTMAIAAQGACR 2220
DB 2161 VTELPVFSRPLQDVVTTEKEKVTLECELSRPNVDVRLKOGVELRAGKTMAIAAQGACR 2220
QY 2221 SLTIYRCFADGVVCDADHAQSSAVKVOQRTVTLIYRRVLADAGEIQFVAENAESR 2280
DB 2221 SLTIYRCFADGVVCDADHAQSSAVKVOQRTVTLIYRRVLADAGEIQFVAENAESR 2280
QY 2281 AOLRVKELPVTILVRPLDKIAMEKHGVLVLECOVSRASAOVRWFKGSQELQPGPKYELVSD 2340
DB 2281 AOLRVKELPVTILVRPLDKIAMEKHGVLVLECOVSRASAOVRWFKGSQELQPGPKYELVSD 2340
QY 2341 GLYRKLIIISDVHAEDDETYTCDAGDVKTSAQPFVEEQSITIVRGLQDVTVMEPAPAWFEC 2400
DB 2341 GLYRKLIIISDVHAEDDETYTCDAGDVKTSAQPFVEEQSITIVRGLQDVTVMEPAPAWFEC 2400
QY 2401 ETSIPSPVPPKWLCKTQLQAGNVGLRQEGTVHRLMLRRTCTMTGTPVHFTVGKSRSSA 2460
DB 2401 ETSIPSPVPPKWLCKTQLQAGNVGLRQEGTVHRLMLRRTCTMTGTPVHFTVGKSRSSA 2460
QY 2461 RLWVSDIIPVLTPLRPLEPTKRELQSVLSCDFRPAPKAVQMYKODTPLSPSEKFKWLSLEG 2520
DB 2461 RLWVSDIIPVLTPLRPLEPTKRELQSVLSCDFRPAPKAVQMYKODTPLSPSEKFKWLSLEG 2520
QY 2521 QMAELRIILRLMPADAGVTRCOAGSAHSSSTVTVAREVTVTGLODABEATBEGWASFCE 2580
DB 2521 QMAELRIILRLMPADAGVTRCOAGSAHSSSTVTVAREVTVTGLODABEATBEGWASFCE 2580
QY 2581 LSHDEEVEWSLNGMPLVNDSPHEISHKGRHRTLVLKSIQRADAGIVRASSIKVSTARS 2640
DB 2581 LSHDEEVEWSLNGMPLVNDSPHEISHKGRHRTLVLKSIQRADAGIVRASSIKVSTARS 2640
QY 2641 EVRVKPVVFLKALDLSAEERTLALQCEVSDPEAHVVRKDGVLQGFSDKYDFLHTAGT 2700
DB 2641 EVRVKPVVFLKALDLSAEERTLALQCEVSDPEAHVVRKDGVLQGFSDKYDFLHTAGT 2700
QY 2701 RGLVVDVSPEDAGLYTCHVGSEETRARVRVHDHVGITKRLKTMVLEBGSCEPVCVLS 2760
DB 2701 RGLVVDVSPEDAGLYTCHVGSEETRARVRVHDHVGITKRLKTMVLEBGSCEPVCVLS 2760
QY 2761 HESASDPAMWTVGGKTVGSSRFQATROGRKYILVVRREAPSDAGEVVVFSVRGLTSKASL 2820
DB 2761 HESASDPAMWTVGGKTVGSSRFQATROGRKYILVVRREAPSDAGEVVVFSVRGLTSKASL 2820
QY 2821 IVRERPAAIIPLEQDQWAPGEDVELRCELSPAGTPVHMLKDKAIRKSQKYDVVCEGTM 2880
DB 2821 IVRERPAAIIPLEQDQWAPGEDVELRCELSPAGTPVHMLKDKAIRKSQKYDVVCEGTM 2880
QY 2881 AMLVIRGASLKDAGEYTCVEASKSTASLHVEEKANCFTTELTNLQVBEKGTAVFTCKTE 2940
DB 2881 AMLVIRGASLKDAGEYTCVEASKSTASLHVEEKANCFTTELTNLQVBEKGTAVFTCKTE 2940
QY 2941 HPAATVTRKGLLELRASGKQPSOEGTLRLTISALEKADSDTYTCDIGQAQSAQLLV 3000
DB 2941 HPAATVTRKGLLELRASGKQPSOEGTLRLTISALEKADSDTYTCDIGQAQSAQLLV 3000
QY 3001 QGRRVHIIEDEDVDVQSGSSATFCRISPNAYEPVHFLDKTPLHANELNEIDAQPGY 3060
DB 3001 QGRRVHIIEDEDVDVQSGSSATFCRISPNAYEPVHFLDKTPLHANELNEIDAQPGY 3060

QY 3061 HVLTLROLALXDSGTIYFEAGDQASAAALRVTEKPSVPSRELTDATITTEGEDTLIVCETS 3120
DB 3061 HVLTLROLALXDSGTIYFEAGDQASAAALRVTEKPSVPSRELTDATITTEGEDTLIVCETS 3120
QY 3121 TCDIPMCWTGDKTILRGSAQCQLSHEGHRAQLLITGATLQDSGRYKCBAGGACSSSIVRV 3180
DB 3121 TCDIPMCWTGDKTILRGSAQCQLSHEGHRAQLLITGATLQDSGRYKCBAGGACSSSIVRV 3180
QY 3181 HARVPRFOEALKOLEVLEGGAAATLRCVLSVAAAPVKWCYGNVLRPGDKYSLRQSGAMLE 3240
DB 3181 HARVPRFOEALKOLEVLEGGAAATLRCVLSVAAAPVKWCYGNVLRPGDKYSLRQSGAMLE 3240
QY 3241 LVVRNLRPODSGRYSCSFGDQTTSATLTVTALPAQFIKLRNKEATEGATATLRCELSKT 3300
DB 3241 LVVRNLRPODSGRYSCSFGDQTTSATLTVTALPAQFIKLRNKEATEGATATLRCELSKT 3300
QY 3301 APVWRKGSSETLRDGDYCLRDQDGMCELOIRGLAMVDAAEYSCVCGEERTSASITIRPM 3360
DB 3301 APVWRKGSSETLRDGDYCLRDQDGMCELOIRGLAMVDAAEYSCVCGEERTSASITIRPM 3360
QY 3361 PAHFIGRLRHQESIEGATATLRCELSKAAPVWRKGRSLRDGDHRSIRQDGAVCLOIC 3420
DB 3361 PAHFIGRLRHQESIEGATATLRCELSKAAPVWRKGRSLRDGDHRSIRQDGAVCLOIC 3420
QY 3421 GLAVADAGEYSCVCGEERTSATLTVKALPAKFTGLRNEEAVEGATAMLCELS KVA PVE 3480
DB 3421 GLAVADAGEYSCVCGEERTSATLTVKALPAKFTGLRNEEAVEGATAMLCELS KVA PVE 3480
QY 3481 WRKGPENLRDGDYTLRQEGTRCELOICGLAMADAGEYLCVCGOERTSATLTIRALPARF 3540
DB 3481 WRKGPENLRDGDYTLRQEGTRCELOICGLAMADAGEYLCVCGOERTSATLTIRALPARF 3540
QY 3541 IEDVKNQDAREGATAVLOCELNSAAPVWRKGSSETLRDGDYSLRQDGTCKELQIRGLAM 3600
DB 3541 IEDVKNQDAREGATAVLOCELNSAAPVWRKGSSETLRDGDYSLRQDGTCKELQIRGLAM 3600
QY 3601 ADTGEYSCVCGOERTSAMLTVRALPIKETEGLRNEEATEGATAVLRCELSKVA PVEWKG 3660
DB 3601 ADTGEYSCVCGOERTSAMLTVRALPIKETEGLRNEEATEGATAVLRCELSKVA PVEWKG 3660
QY 3661 HETLRDGRHSLRDQDGMCELOIRGLVAEDAGEYLCMCKERTSAMLTVRAMPSPFIEGL 3720
DB 3661 HETLRDGRHSLRDQDGMCELOIRGLVAEDAGEYLCMCKERTSAMLTVRAMPSPFIEGL 3720
QY 3721 RNEEATEGDTATLWCELSKAAAPVWRKGHETLRDGRHSLRQDGRSCELOIRGLAVVDAG 3780
DB 3721 RNEEATEGDTATLWCELSKAAAPVWRKGHETLRDGRHSLRQDGRSCELOIRGLAVVDAG 3780
QY 3781 EYSCVCGOERTSATLTVRALPARFIEDVKNQDAREGATAVLOCELNSKAA PVEWKGSETL 3840
DB 3781 EYSCVCGOERTSATLTVRALPARFIEDVKNQDAREGATAVLOCELNSKAA PVEWKGSETL 3840
QY 3841 RGGDRYSLRQDGTTRCELOIHGLSVADTGEYSCVCGOERTSATLTVRAPQVPRPRELQSLQ 3900
DB 3841 RGGDRYSLRQDGTTRCELOIHGLSVADTGEYSCVCGOERTSATLTVRAPQVPRPRELQSLQ 3900
QY 3901 AEEGSTATLQCELSEPTATVVMKGGLOLQANGRRPRELQGTAEVLQDLOREDTGEYT 3960
DB 3901 AEEGSTATLQCELSEPTATVVMKGGLOLQANGRRPRELQGTAEVLQDLOREDTGEYT 3960
QY 3961 CTCGSOATSATLTVTAAAPVRLRELOHQVDEGGTAHLCCELSRAGASVWRKGSLOLPP 4020
DB 3961 CTCGSOATSATLTVTAAAPVRLRELOHQVDEGGTAHLCCELSRAGASVWRKGSLOLPP 4020
QY 4021 CAKYQWODGAAELLVRGVEQEDAGDYTCDTGHTQSMASLSVRVPRPKFKTRLSQLEQ 4080
DB 4021 CAKYQWODGAAELLVRGVEQEDAGDYTCDTGHTQSMASLSVRVPRPKFKTRLSQLEQ 4080
QY 4081 TGDILARLCCQLSDAESGAVVQWMLKEGVELHAGPKYEMRSQATRELLIHOLEAKDTGEYA 4140
DB 4081 TGDILARLCCQLSDAESGAVVQWMLKEGVELHAGPKYEMRSQATRELLIHOLEAKDTGEYA 4140
QY 4141 CVTGGQKTAASLRVTEPEVTIVRGLVDAEVTADEDEVEFSCEVSRAGATGVQWCLQGLPIQ 4200

Db 4141 CVTGGOKTAAASLRVTEPEVTI VRGLVDAEVTADEDEVEFSCEVSRAGATGVQWCLQGLPLQ 4200
QY SNEVTEVAVRDGRIHTLRKLVGTPEDAGTVSHLGNHASSAOLTVRAPEVTTILEPLODVQ 4260
Db SNEVTEVAVRDGRIHTLRKLVGTPEDAGTVSHLGNHASSAOLTVRAPEVTTILEPLODVQ 4260
QY LSEGQDASFQCLRSRASGOEARWALGGVPLQANEMNDITVEQOTLHLTLHKVTTLEDAGT 4320
Db LSEGQDASFQCLRSRASGOEARWALGGVPLQANEMNDITVEQOTLHLTLHKVTTLEDAGT 4320
QY VSFHVGTCSSEAOQLKVTAKNTVVRGLENVAEALBEGGALFECOLSOPAEVAHWTLLDDBPV 4380
Db VSFHVGTCSSEAOQLKVTAKNTVVRGLENVAEALBEGGALFECOLSOPAEVAHWTLLDDBPV 4380
QY RTSSENAVVPFFENGLRHLLKLNLRPODSCRVTFLAGDMVTSAFITVRGMRLEIILEPLKN 4440
Db RTSSENAVVPFFENGLRHLLKLNLRPODSCRVTFLAGDMVTSAFITVRGMRLEIILEPLKN 4440
QY AAVRAGAQAARFTCTLSEAVPVGEASWYINGAAVQPDSDWTVTADGSHQALLRSAQPHH 4500
Db AAVRAGAQAARFTCTLSEAVPVGEASWYINGAAVQPDSDWTVTADGSHQALLRSAQPHH 4500
QY AGEVTFACHDVAVASARLTVLGLPDPPEDEAVVAHSHSTVTLISWAAPMGSDGGGLCGYRVE 4560
Db AGEVTFACHDVAVASARLTVLGLPDPPEDEAVVAHSHSTVTLISWAAPMGSDGGGLCGYRVE 4560
QY VKEGATGQWRLCHELVPGPECVVDGLA PGETVFRVAAVGPVGAPEVHLPTQVRLAEP 4620
Db VKEGATGQWRLCHELVPGPECVVDGLA PGETVFRVAAVGPVGAPEVHLPTQVRLAEP 4620
QY KVPVPPQSPAPESRQVAAGSDVLSLEVAEAGEV IWHKGMER IOPGGRFEVVSQGRQQL 4680
Db KVPVPPQSPAPESRQVAAGSDVLSLEVAEAGEV IWHKGMER IOPGGRFEVVSQGRQQL 4680
QY VTKGFTAEDQGVHCGLAGQSCIPAAATFOVALSPASVDEAPOPSILPPEAAQEGDLHLI 4740
Db VTKGFTAEDQGVHCGLAGQSCIPAAATFOVALSPASVDEAPOPSILPPEAAQEGDLHLI 4740
QY BALARKRMSREPTLDSISELPEEDGRSQRLOPQAEAEVAPDLSEGYSTADELARTGDADL 4800
Db BALARKRMSREPTLDSISELPEEDGRSQRLOPQAEAEVAPDLSEGYSTADELARTGDADL 4800
QY SHTSDDSDSRAGTSPSLVTLKAGRPGTSPASKVGAPAAPSVKPPQOOEPLAAVRPLG 4860
Db SHTSDDSDSRAGTSPSLVTLKAGRPGTSPASKVGAPAAPSVKPPQOOEPLAAVRPLG 4860
QY DLSTKDLGDPMSDKAAVKIQAAFKGYKVRKEMKQOEGPWFSHRTGDTAEAOVGDALRLCEV 4920
Db DLSTKDLGDPMSDKAAVKIQAAFKGYKVRKEMKQOEGPWFSHRTGDTAEAOVGDALRLCEV 4920
QY VASKADVRARWLKDGVELTDGRHHIIDLQGDGTCSLIAGLDRADAGCYTCQVSNKFGQV 4980
Db VASKADVRARWLKDGVELTDGRHHIIDLQGDGTCSLIAGLDRADAGCYTCQVSNKFGQV 4980
QY THSACVVVSGSEASESSGGELDADFRRAARRLHRLPRTKSPAESVDEELFLSADEGPA 5040
Db THSACVVVSGSEASESSGGELDADFRRAARRLHRLPRTKSPAESVDEELFLSADEGPA 5040
QY EPEEPADQVTYREDHFI CIRFEALTEARQAVTRFOEMFATLGIGVRIKLVQEQPRVEM 5100
Db EPEEPADQVTYREDHFI CIRFEALTEARQAVTRFOEMFATLGIGVRIKLVQEQPRVEM 5100
QY CISKETPAPVPPPEPLPSLLTSDAAPVFLTELQNEQVQDGYPVSDCVVTGQPMPSVRWF 5160
Db CISKETPAPVPPPEPLPSLLTSDAAPVFLTELQNEQVQDGYPVSDCVVTGQPMPSVRWF 5160
QY KQGLLEDDHYMINEOQGHQI IITAVNPADMGVVRCLAENMGVSSKAEILRVDLTS 5220
Db KQGLLEDDHYMINEOQGHQI IITAVNPADMGVVRCLAENMGVSSKAEILRVDLTS 5220
QY TDYDTAADATBESSYFSAQGYLSSREOGBTSTTDEGOLPOVBEELRLDQVAPGTRLAKF 5280

Db 5221 TDYDTAADATBESSYFSAQGYLSSREOGBTSTTDEGOLPOVBEELRLDQVAPGTRLAKF 5280
QY OLKVGYPAPRLYMFKQGOPLTASAHIRMTCKKILHTLEIISVTREDSGQVAAIISNANG 5340
Db OLKVGYPAPRLYMFKQGOPLTASAHIRMTCKKILHTLEIISVTREDSGQVAAIISNANG 5340
QY AAYSARLLVRGPDPEPEKPSADVHEOLVPPRMLERFTPKKVKKSSSTTFPSVKVGEGRVVP 5400
Db AAYSARLLVRGPDPEPEKPSADVHEOLVPPRMLERFTPKKVKKSSSTTFPSVKVGEGRVVP 5400
QY TVHMLREBEARGVLIWIGPDTGYTVASSAQHSLVLLDVGROHQGTVTCIASNAAGQALC 5460
Db TVHMLREBEARGVLIWIGPDTGYTVASSAQHSLVLLDVGROHQGTVTCIASNAAGQALC 5460
QY SASLHVSGLPKVEQEKVKEALISTFLOGTTOAISAOGLETTASADLGGOKKEEPLAKE 5520
Db SASLHVSGLPKVEQEKVKEALISTFLOGTTOAISAOGLETTASADLGGOKKEEPLAKE 5520
QY ALGHLSLAEVGTETEFLOKLTSQITEMVSAKITQAKLOVPGGSDSDSKTSPASPRHGRSR 5580
Db ALGHLSLAEVGTETEFLOKLTSQITEMVSAKITQAKLOVPGGSDSDSKTSPASPRHGRSR 5580
QY PSSSIQSSSESESDGDARGEIFDIYVVTADYLP LGAEQDAITTLREGQYVEVLDAAHPLRW 5640
Db PSSSIQSSSESESDGDARGEIFDIYVVTADYLP LGAEQDAITTLREGQYVEVLDAAHPLRW 5640
QY LVRTKPTKSSPSRQGWSPAYLDRRLKLSPEWGAABAPPEPGEAVSEDEYKARLSSVIOE 5700
Db LVRTKPTKSSPSRQGWSPAYLDRRLKLSPEWGAABAPPEPGEAVSEDEYKARLSSVIOE 5700
QY LLSSEQAQFVBELOQLQSHHLQHLERCPHPVPIAVAGQKAVIFRNVRTDGRPHSSFLQBLQQ 5760
Db LLSSEQAQFVBELOQLQSHHLQHLERCPHPVPIAVAGQKAVIFRNVRTDGRPHSSFLQBLQQ 5760
QY CDTDDDDVAMCFIKNOAFAEQYLEFLVGRVOAESVWVSTALQEFYKKYKVAEBALLAGDPSQP 5820
Db CDTDDDDVAMCFIKNOAFAEQYLEFLVGRVOAESVWVSTALQEFYKKYKVAEBALLAGDPSQP 5820
QY PPPPLQHYLQBPVERVORYQALLKELIRNKARNQNCALLQEAQAVVVSALLPQAEAKLHV 5880
Db PPPPLQHYLQBPVERVORYQALLKELIRNKARNQNCALLQEAQAVVVSALLPQAEAKLHV 5880
QY SLMENYPTGLLEALGEPITROGHFIWEGAPGARMPWKGNHRHFLFRNHLVICPRDSDRT 5940
Db SLMENYPTGLLEALGEPITROGHFIWEGAPGARMPWKGNHRHFLFRNHLVICPRDSDRT 5940
QY DTYSYVFRNMKLLSSIDLNDQVEGDDRAFPWQEREDSVRKYLLOARTAIKSSWVKEIC 6000
Db DTYSYVFRNMKLLSSIDLNDQVEGDDRAFPWQEREDSVRKYLLOARTAIKSSWVKEIC 6000
QY GIQORLALPVRRPPDFEEELADCTAELGETVKLACRVGTGPKPVIWYKDGKAVQVDPHH 6060
Db GIQORLALPVRRPPDFEEELADCTAELGETVKLACRVGTGPKPVIWYKDGKAVQVDPHH 6060
QY ILIEDPDGSCALILDSLTGVDSQYMCFAASAAGNCSTLGKILVQVPPFRFNKVRASPFV 6120
Db ILIEDPDGSCALILDSLTGVDSQYMCFAASAAGNCSTLGKILVQVPPFRFNKVRASPFV 6120
QY EGEDAQFTCTTIEGAPYQIRWYKDGALLTTGNKFTLSEPRSGLLVIVIRAASKEDLGLY 6180
Db EGEDAQFTCTTIEGAPYQIRWYKDGALLTTGNKFTLSEPRSGLLVIVIRAASKEDLGLY 6180
QY ECELNVRLGASARASAEIRIQSPMLQAEQCHREQLVAAVEVTEQETKVPKKTVIISETIT 6240
Db ECELNVRLGASARASAEIRIQSPMLQAEQCHREQLVAAVEVTEQETKVPKKTVIISETIT 6240
QY SVLKRLLLGPKAPGSPSTGDLTGGPCPRG-----APAL-----QETSGQP- 6271
Db SVLKRLLLGPKAPGSPSTGDLTGGPCPRG-----APAL-----QETSGQP- 6271
QY TVVKSPPRQRR--SPSKSPSRSPSRCSASPLRPGLLAPDLILYLPAGCQPPRRPEAPQKPV 6299
Db TVVKSPPRQRR--SPSKSPSRSPSRCSASPLRPGLLAPDLILYLPAGCQPPRRPEAPQKPV 6299
QY -PVYGTSEAPVPPRV- ---POPILLHGEPE-----QEPETAIARAOEWTPVIRMEGA 6317
Db -PVYGTSEAPVPPRV- ---POPILLHGEPE-----QEPETAIARAOEWTPVIRMEGA 6317
QY VPTLYYTEANISPALPGLSGFPQKPVKEVEETIEVVRKMGKQGVSPFTTE--VPRSSSGH 6357
Db VPTLYYTEANISPALPGLSGFPQKPVKEVEETIEVVRKMGKQGVSPFTTE--VPRSSSGH 6357

QY 6318 AW--PGCTG 6325
DB 6358 LFTLPGATPG 6367

RESULT 3

US-10-408-765A-2291
; Sequence 2291, Application US/10408765A
; Publication No. US20040101874A1
; GENERAL INFORMATION:
; APPLICANT: Ghosh, Soumitra S.
; APPLICANT: Faby, Boin D.
; APPLICANT: Zhang, Bing
; APPLICANT: Gibson, Bradford W.
; APPLICANT: Taylor, Steven W.
; APPLICANT: Glenn, Gary M.
; APPLICANT: Warnock, Dale E.
; TITLE OF INVENTION: TARGETS FOR THERAPEUTIC INTERVENTION
; TITLE OF INVENTION: IDENTIFIED IN THE MITOCHONDRIAL PROTEOME
; FILE REFERENCE: 66088-465
; CURRENT APPLICATION NUMBER: US/10/408,765A
; CURRENT FILING DATE: 2003-04-04
; NUMBER OF SEQ ID NOS: 3077
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2291
; LENGTH: 6620
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-408-765A-2291

Query Match 77.9%; Score 32134; DB 16; Length 6620;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 6250; Conservative 15; Mismatches 57; Indels 48; Gaps 9;
QY 1 MDQPFSGAPRLTPKAPVSVGKDATLSQIVGNPTPOVSWEXDQOPVTAGARFLAQ 60
DB 1 MDQPFSGAPRLTPKAPVSVGKDATLSQIVGNPTPOVSWEXDQOPVTAGARFLAQ 60
QY 61 DGDYRLTLDLALGDSGOYVCRARNAIGEAFVAVGLQVDAEAAEQAHPFLRPTSI 120
DB 61 DGDYRLTLDLALGDSGOYVCRARNAIGEAFVAVGLQVDAEAAEQAHPFLRPTSI 120
QY 121 VREGSEATFRCVGSGPRPAVSWSDGRRLGPDGPRVVRBELGEASALIRAAAPRDGG 180
DB 121 VREGSEATFRCVGSGPRPAVSWSDGRRLGPDGPRVVRBELGEASALIRAAAPRDGG 180
QY 181 TYEVRANPLGAASAAALVDSDAADTASRPGTSTAALLAHLORREARAEAGAPASP 240
DB 181 TYEVRANPLGAASAAALVDSDAADTASRPGTSTAALLAHLORREARAEAGAPASP 240
QY 241 STGTRTCTVTGEGHARLSCTYVTGEPKPTVMKKDQQLVTEGRRHVYVEDAQENFVLKILF 300
DB 241 STGTRTCTVTGEGHARLSCTYVTGEPKPTVMKKDQQLVTEGRRHVYVEDAQENFVLKILF 300
QY 301 CKQSDRGILYCTASNLVQGTYSVLVVRPAVPPKQLDLVREKESATFLCEVPQPS 360
DB 301 CKQSDRGILYCTASNLVQGTYSVLVVRPAVPPKQLDLVREKESATFLCEVPQPS 360
QY 361 TEAAWFKETRLWASAKYIGIEEGTERLTVRNVSAADDVYICETPGSRVVAELAVQG 420
DB 361 TEAAWFKETRLWASAKYIGIEEGTERLTVRNVSAADDVYICETPGSRVVAELAVQG 420
QY 421 NLLRLPKRTAVRVDGTAMFCVELAVPGPVHWRNQEVEVAGGRVAISAEGTRHTLTIS 480
DB 421 NLLRLPKRTAVRVDGTAMFCVELAVPGPVHWRNQEVEVAGGRVAISAEGTRHTLTIS 480
QY 481 QCCLEDVQVAFMAGDCOTSTRFCVSAAPRKPLQPPDPVVKARMESSVILSWSPPHGE 540
DB 481 QCCLEDVQVAFMAGDCOTSTRFCVSAAPRKPLQPPDPVVKARMESSVILSWSPPHGE 540
QY 541 RPVTIDGYLVKCKLGTYYTWIRCHAEAWVATPELTADVABEGNFQFRVSAALNSFGQSPY 600
DB 541 RPVTIDGYLVKCKLGTYYTWIRCHAEAWVATPELTADVABEGNFQFRVSAALNSFGQSPY 600

DB 541 RPVTIDGYLVKCKLGTYYTWIRCHAEAWVATPELTADVABEGNFQFRVSAALNSFGQSPY 600
QY 601 LEFGPTVHLAPKLAVRTPLKAVQAVEGGEVTFPSVDLTVASAGFWLQALKASSVYETH 660
DB 601 LEFGPTVHLAPKLAVRTPLKAVQAVEGGEVTFPSVDLTVASAGFWLQALKASSVYETH 660
QY 661 CDTRHTLTITREVPAHLGAOLKVFVANGIESSIRNEVRAAPGLTANKPPAAAAREVLARL 720
DB 661 CDTRHTLTITREVPAHLGAOLKVFVANGIESSIRNEVRAAPGLTANKPPAAAAREVLARL 720
QY 721 HEEAQLLAELSDQAAAVTWLKDGRITLSPGPKEVQASAGRRVLLVRDVARDDAGLYECVS 780
DB 721 HEEAQLLAELSDQAAAVTWLKDGRITLSPGPKEVQASAGRRVLLVRDVARDDAGLYECVS 780
QY 781 RGGRIAYQLSVQGLARFLHKDMAGSCVDVAGGPAQFECETSEAHVHVHWYKDGMELGHS 840
DB 781 RGGRIAYQLSVQGLARFLHKDMAGSCVDVAGGPAQFECETSEAHVHVHWYKDGMELGHS 840
QY 841 GERFLOEDVGTTRHLVAATVTRDEGTYSCVGEDSVDFRLRVSEPKVVPFAKEQLARRKL 900
DB 841 GERFLOEDVGTTRHLVAATVTRDEGTYSCVGEDSVDFRLRVSEPKVVPFAKEQLARRKL 900
QY 901 QAEAGASATLSCEVAQAQTEVTWYKDGKLSKSSSKVCMCATGCTRRLLVVQAGQADAGEY 960
DB 901 QAEAGASATLSCEVAQAQTEVTWYKDGKLSKSSSKVCMCATGCTRRLLVVQAGQADAGEY 960
QY 961 SCEAGGQRLSPHLDVKEPKVVFAXQVAHSEVQAEAGANATLSCEVAQAQAEVMMYKDGK 1020
DB 961 SCEAGGQRLSPHLDVKEPKVVFAXQVAHSEVQAEAGANATLSCEVAQAQAEVMMYKDGK 1020
QY 1021 KLSSSLKVHVEAKGCRRLVVQAGTKTDAGDYSCEARGQVSRFLHITTEPKMFAKEQSV 1080
DB 1021 KLSSSLKVHVEAKGCRRLVVQAGTKTDAGDYSCEARGQVSRFLHITTEPKMFAKEQSV 1080
QY 1081 HNEVQAEAGASAMLSCEVAQAQTEVTWYKDGKLSKSSSKVMEVKGCTRRLLVPOAGKAD 1140
DB 1081 HNEVQAEAGASAMLSCEVAQAQTEVTWYKDGKLSKSSSKVMEVKGCTRRLLVPOAGKAD 1140
QY 1141 AGEYSCEAGGQVRSFHLHITTEPKGVFAKEQSVHNEVQAEAGTAMLSCEVAQAQTEVTWY 1200
DB 1141 AGEYSCEAGGQVRSFHLHITTEPKGVFAKEQSVHNEVQAEAGTAMLSCEVAQAQTEVTWY 1200
QY 1201 KDGKLSKSSSKVMEVKGCTRRLLVVQVQKADAGEYSCEAGGQVRSFOLHITTEPKAVPAK 1260
DB 1201 KDGKLSKSSSKVMEVKGCTRRLLVVQVQKADAGEYSCEAGGQVRSFOLHITTEPKAVPAK 1260
QY 1261 EQLVHNEVTRTEAGASATLSCEVAQAQTEVTWYKDGKLSKSSSKVRIEAGCMRQLVVOQA 1320
DB 1261 EQLVHNEVTRTEAGASATLSCEVAQAQTEVTWYKDGKLSKSSSKVRIEAGCMRQLVVOQA 1320
QY 1321 QADAGEYTCBAGGQRLSPHLDVSEPKAVFAKEQLAHRKVQAEAGAIATLSCEVAQAQTE 1380
DB 1321 QADAGEYTCBAGGQRLSPHLDVSEPKAVFAKEQLAHRKVQAEAGAIATLSCEVAQAQTE 1380
QY 1381 VTWYKDGKLSKSSSKVMEVKGCTRRLLVVQQAQADAGEYSCEAGGQRLSPSLDVAEPKV 1440
DB 1381 VTWYKDGKLSKSSSKVMEVKGCTRRLLVVQQAQADAGEYSCEAGGQRLSPSLDVAEPKV 1440
QY 1441 VPAKEQPVHREVOQAQAGSTTSLCEVAQAQTEVTWYKDGKLSKSSSKVMEVKGCTRRLLV 1500
DB 1441 VPAKEQPVHREVOQAQAGSTTSLCEVAQAQTEVTWYKDGKLSKSSSKVMEVKGCTRRLLV 1500
QY 1501 VQQAQADAGEYSCEAGGQRLSPHLDVSEPKAVFAKEQVPAKFAKEQVPAKFAKEQVPAKFAKEQV 1560
DB 1501 VQQAQADAGEYSCEAGGQRLSPHLDVSEPKAVFAKEQVPAKFAKEQVPAKFAKEQVPAKFAKEQV 1560
QY 1561 AQTEVTWYKDGKLSKSSSKVMEVKGCTRRLLVVQQAQADAGEYSCKAGDQRLSPHLDVA 1620
DB 1561 AQTEVTWYKDGKLSKSSSKVMEVKGCTRRLLVVQQAQADAGEYSCKAGDQRLSPHLDVA 1620
QY 1621 EPKVVFAKEQVPAHREVOQAAGASATLSCEVAQAQTEVTWYKDGKLSKSSSKVMEVKGCT 1680
DB 1621 EPKVVFAKEQVPAHREVOQAAGASATLSCEVAQAQTEVTWYKDGKLSKSSSKVMEVKGCT 1680

Qy	1681	RRLVVQAGQADAGEYCEAGQORLSRLHVAELEPOISERPCHREPLVWKEHEDIILTA	1740	Qy	2761	HESASDPAMWTVGKTVGSSSRFOATROGRKYLIVVREAAAPSDAGEVVFVSRGLTYSKASL	2820
Db	1681	RRLVVQAGQADAGEYCEAGQORLSRLHVAELEPOISERPCHREPLVWKEHEDIILTA	1740	Db	2761	HESASDPAMWTVGKTVGSSSRFOATROGRKYLIVVREAAAPSDAGEVVFVSRGLTYSKASL	2820
Qy	1741	TLATPSAATVWLKDGVEIRRSKHETASQDTHTLTVHGAQVLDLSAIYSRCVGAEGQDF	1800	Qy	2821	IYVERPAAIIPKPLEDQWVARGEDVELCELSRAGTPVHMLKDKRAIRKSOKYDVVCEGTM	2880
Db	1741	TLATPSAATVWLKDGVEIRRSKHETASQDTHTLTVHGAQVLDLSAIYSRCVGAEGQDF	1800	Db	2821	IYVERPAAIIPKPLEDQWVARGEDVELCELSRAGTPVHMLKDKRAIRKSOKYDVVCEGTM	2880
Qy	1801	PVQVEEVAAPCRLLPEVGBGLGTVTLACBLSAPACAEVWRCNGTQPRVKRQFQMAEG	1860	Qy	2881	AMLVIRGASLKDAGEYTCVEASKSTASLHVEBKANCFTBELTNLQVEEKGTAFTVCTKTE	2940
Db	1801	PVQVEEVAAPCRLLPEVGBGLGTVTLACBLSAPACAEVWRCNGTQPRVKRQFQMAEG	1860	Db	2881	AMLVIRGASLKDAGEYTCVEASKSTASLHVEBKANCFTBELTNLQVEEKGTAFTVCTKTE	2940
Qy	1861	PVRSITVLGLRAEDAGEVVCESRDHDSAQITVSPRVVKFMSGSLTVVAEEGGEATFOC	1920	Qy	2941	HPAATVTRWKGLELLELRASGHQPSQEGTLRLTISALEKADSDTYTCDIQOASRAQLLV	3000
Db	1861	PVRSITVLGLRAEDAGEVVCESRDHDSAQITVSPRVVKFMSGSLTVVAEEGGEATFOC	1920	Db	2941	HPAATVTRWKGLELLELRASGHQPSQEGTLRLTISALEKADSDTYTCDIQOASRAQLLV	3000
Qy	1921	VVSPSDVAVWFRDGLLQPSSEKFAISQSGASHSLTISDLVLEDAQOITVBAEGASSAA	1980	Qy	3001	QGRVHIIEDLEDDVDVQEGSSATFRCHRI SPANTYEPVHWFLLDKTPLHANELNEIDAQPGGY	3060
Db	1921	VVSPSDVAVWFRDGLLQPSSEKFAISQSGASHSLTISDLVLEDAQOITVBAEGASSAA	1980	Db	3001	QGRVHIIEDLEDDVDVQEGSSATFRCHRI SPANTYEPVHWFLLDKTPLHANELNEIDAQPGGY	3060
Qy	1981	LRVREAPVLFKKLEPQTVBERSSVTLVELTRPWPPELRWTNATNAPGKNVEIHAEGA	2040	Qy	3061	HVLTLRQALALKDSTIYFEAGDQASAAALRVTEKPSVFSRELTDATITTEGEDTLVLCETS	3120
Db	1981	LRVREAPVLFKKLEPQTVBERSSVTLVELTRPWPPELRWTNATNAPGKNVEIHAEGA	2040	Db	3061	HVLTLRQALALKDSTIYFEAGDQASAAALRVTEKPSVFSRELTDATITTEGEDTLVLCETS	3120
Qy	2041	RHRLVLHNVGFADRGFFCETPDDKTQAKLTVMRQVRLVRGLQAVEAREOGTATMEVOL	2100	Qy	3121	TCDIPMCWTQDKTLRGASARCOLSHEGHRAQLLITGATLQDSGRYKCEAGCACSSIVRV	3180
Db	2041	RHRLVLHNVGFADRGFFCETPDDKTQAKLTVMRQVRLVRGLQAVEAREOGTATMEVOL	2100	Db	3121	TCDIPMCWTQDKTLRGASARCOLSHEGHRAQLLITGATLQDSGRYKCEAGCACSSIVRV	3180
Qy	2101	SHADVDGSWTRDGLRFQOGPTCHLAVRGPMTLTLISGLRPEDSGLMVFKASGVHTSARLV	2160	Qy	3181	HARVPRFOEALKOLEVLEGGAAUCLRVLSSVAAAPVKYGNVLRPGDKYSLROEGAMLE	3240
Db	2101	SHADVDGSWTRDGLRFQOGPTCHLAVRGPMTLTLISGLRPEDSGLMVFKASGVHTSARLV	2160	Db	3181	HARVPRFOEALKOLEVLEGGAAUCLRVLSSVAAAPVKYGNVLRPGDKYSLROEGAMLE	3240
Qy	2161	VTELPVSRPLQDVVTTKEKVTLECELSRPNDVRLKDGVELRAGKTWALIAAQGACR	2220	Qy	3241	LVVNLRLPQDSGRYSCSFGDQTTSATLTVTALPAQFTGKLRNKEATEGATLRLCELSKT	3300
Db	2161	VTELPVSRPLQDVVTTKEKVTLECELSRPNDVRLKDGVELRAGKTWALIAAQGACR	2220	Db	3241	LVVNLRLPQDSGRYSCSFGDQTTSATLTVTALPAQFTGKLRNKEATEGATLRLCELSKT	3300
Qy	2221	SITIYRCEFPADQGVVCDADHAQSSASVKVQGRTYTLIYRVLAEADAGEIQFVAENAESR	2280	Qy	3301	APVEMWKGSETLRDGDYCLRDQAMGCELOIRGLAMVDAAVYSCVCGEERTSASITIRPM	3360
Db	2221	SITIYRCEFPADQGVVCDADHAQSSASVKVQGRTYTLIYRVLAEADAGEIQFVAENAESR	2280	Db	3301	APVEMWKGSETLRDGDYCLRDQAMGCELOIRGLAMVDAAVYSCVCGEERTSASITIRPM	3360
Qy	2281	AQLRKVELPVTILVRPLRDKIAEMKRGVLECOVSRASAOVRWFKGSQELQPGPKYELVSD	2340	Qy	3361	PAHFIGRLRHQESLEGATATLRLCELSKAAAPVEMWKGRESLRDGDHSLRLQDGA VCELQIC	3420
Db	2281	AQLRKVELPVTILVRPLRDKIAEMKRGVLECOVSRASAOVRWFKGSQELQPGPKYELVSD	2340	Db	3361	PAHFIGRLRHQESLEGATATLRLCELSKAAAPVEMWKGRESLRDGDHSLRLQDGA VCELQIC	3420
Qy	2341	GLYRKLIIISDVHAEDEDTYCDAGDVKTSAQPFVEEQSITIVRGLODVTVMPEAPAWPEC	2400	Qy	3421	GLAVADAGEYSCVCGEERTSATLTVKALPAKFTGLRNEEAVEGATAMLWCELSKVAPE	3480
Db	2341	GLYRKLIIISDVHAEDEDTYCDAGDVKTSAQPFVEEQSITIVRGLODVTVMPEAPAWPEC	2400	Db	3421	GLAVADAGEYSCVCGEERTSATLTVKALPAKFTGLRNEEAVEGATAMLWCELSKVAPE	3480
Qy	2401	ETSIPIPSVRPPKWLKGTVLQAGNVGLEQEGTVHRLMLRRCTSTMTGPHFTVVKSRSSA	2460	Qy	3481	WRKGPENLRDGDYILRQEGTRCELOICGLAMADAGEYLCVCGOERTSATLITRALPARF	3540
Db	2401	ETSIPIPSVRPPKWLKGTVLQAGNVGLEQEGTVHRLMLRRCTSTMTGPHFTVVKSRSSA	2460	Db	3481	WRKGPENLRDGDYILRQEGTRCELOICGLAMADAGEYLCVCGOERTSATLITRALPARF	3540
Qy	2461	RLUVSDIIPVLTREPLETKGRELQSVLSCDPRPAPKAVQWYKDDTPLSPSEKFKWSLEG	2520	Qy	3541	IEDVKNOEAREGATAVLQCELSNAAAPVEMWKGSETLRDGDYRSLRQDGTCKEQLIRGLAM	3600
Db	2461	RLUVSDIIPVLTREPLETKGRELQSVLSCDPRPAPKAVQWYKDDTPLSPSEKFKWSLEG	2520	Db	3541	IEDVKNOEAREGATAVLQCELSNAAAPVEMWKGSETLRDGDYRSLRQDGTCKEQLIRGLAM	3600
Qy	2521	QMAELRIILMPADAGVYRCQAGSAHSSTEVTVEAREVTVTGLODAEATEEGWASFCE	2580	Qy	3601	ADTGEYSCVCGOERTSATLTVRALPIKFTEGLRNEEATEGATAVLRCELSKMAPEVWKG	3660
Db	2521	QMAELRIILMPADAGVYRCQAGSAHSSTEVTVEAREVTVTGLODAEATEEGWASFCE	2580	Db	3601	ADTGEYSCVCGOERTSATLTVRALPIKFTEGLRNEEATEGATAVLRCELSKMAPEVWKG	3660
Qy	2581	LSHEDEEVEWSLNGMPLYNDSFHEISHKGRHTLVLSKTORADAGIVRASSILKVSTARS	2640	Qy	3661	HETLRDGDHSLRQDGA RCELOIRGLVAEDAGEYLCMCGKERTSAMLTVRAMPKSFIEGL	3720
Db	2581	LSHEDEEVEWSLNGMPLYNDSFHEISHKGRHTLVLSKTORADAGIVRASSILKVSTARS	2640	Db	3661	HETLRDGDHSLRQDGA RCELOIRGLVAEDAGEYLCMCGKERTSAMLTVRAMPKSFIEGL	3720
Qy	2641	EVVRKPVVFLKALDDLAEERGTALQCEVSDPEAHVWVRKQVQGPSDKDYDLHTAGT	2700	Qy	3721	RNEEATEGDTATLWCELSKAAAPVEMWKGSETLRDGDHSLRQDGSRCCELOIRGLAVDAG	3780
Db	2641	EVVRKPVVFLKALDDLAEERGTALQCEVSDPEAHVWVRKQVQGPSDKDYDLHTAGT	2700	Db	3721	RNEEATEGDTATLWCELSKAAAPVEMWKGSETLRDGDHSLRQDGSRCCELOIRGLAVDAG	3780
Qy	2701	RLGVHDVSPEDAGLYTCHVGSEETRARVRVHDLHVGITKRLKTMTEVLGEGSCSFECVL	2760	Qy	3781	EYSVCGOERTSATLTVRALPARFIEDVKNOEAREGATAVLQCELSKAAPEVWKGSETL	3840
Db	2701	RLGVHDVSPEDAGLYTCHVGSEETRARVRVHDLHVGITKRLKTMTEVLGEGSCSFECVL	2760	Db	3781	EYSVCGOERTSATLTVRALPARFIEDVKNOEAREGATAVLQCELSKAAPEVWKGSETL	3840
				Qy	3841	RGGDRYSLRQDGT RCELOIRGLHGLSVADTGEYSVCGOERTSATLTVRAPQVFPVFPQSLQ	3900

Db	3841	RGDRYSLRQDGTGRCBLQIHGLSVADTGEYSVCVGBRTSATLTVRAPOQVFPREPLQSLQ	3900
Qy	3901	ABEGSTATLOCELSEPTATVWVSKGGLQLOANGRRPRLOCTAELVLQLOREDTGEYT	3960
Db	3901	ABEGSTATLOCELSEPTATVWVSKGGLQLOANGRRPRLOCTAELVLQLOREDTGEYT	3960
Qy	3961	CTCGSQATSLTVTAAPVRFLRELOHQEVDGCTAHLCCCELGRAGASVEWRKGSLOLFP	4020
Db	3961	CTCGSQATSLTVTAAPVRFLRELOHQEVDGCTAHLCCCELGRAGASVEWRKGSLOLFP	4020
Qy	4021	CAKYQWQDGAABELLVRGVEQEDAGDYTCDTGHTQSMASLSVRVPRPKTKTLOSLEQ	4080
Db	4021	CAKYQWQDGAABELLVRGVEQEDAGDYTCDTGHTQSMASLSVRVPRPKTKTLOSLEQ	4080
Qy	4081	TGDIARLCCOLSDAESGAVVQWLKEGVHELHAGPKYEMRSOGATRELLIHQLEAKDTGEYA	4140
Db	4081	TGDIARLCCOLSDAESGAVVQWLKEGVHELHAGPKYEMRSOGATRELLIHQLEAKDTGEYA	4140
Qy	4141	CVTGGQKTAASLRVTEPEVTIVRGLVDAEVTADDEVEFSCEVSRAGATGVQWCLQGLPLQ	4200
Db	4141	CVTGGQKTAASLRVTEPEVTIVRGLVDAEVTADDEVEFSCEVSRAGATGVQWCLQGLPLQ	4200
Qy	4201	SNEVTEVAVRDGRHTLRLKGVTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPQDVQ	4260
Db	4201	SNEVTEVAVRDGRHTLRLKGVTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPQDVQ	4260
Qy	4261	LSEQODASFOCRLSRASQGEARWALGVPLOANEMNDITVEQGTGLHLLTLHKVTLDEAGT	4320
Db	4261	LSEQODASFOCRLSRASQGEARWALGVPLOANEMNDITVEQGTGLHLLTLHKVTLDEAGT	4320
Qy	4321	VSPHVGTCSSEAQLKVTAKNTVVRGLNVLEALGGEALFECQLSQPEVAHTWLLDPEV	4380
Db	4321	VSPHVGTCSSEAQLKVTAKNTVVRGLNVLEALGGEALFECQLSQPEVAHTWLLDPEV	4380
Qy	4381	RTSENABVVPFENGLRHLHLLKNLRPODCSRVTFPLAGDMVTSAPLTVRGWRLEIILEPLKN	4440
Db	4381	RTSENABVVPFENGLRHLHLLKNLRPODCSRVTFPLAGDMVTSAPLTVRGWRLEIILEPLKN	4440
Qy	4441	AAVRAGAQAARFTCTLSEAVPGEASWYINGAAVQDDSDMTVTADGSHQALLLRSQAOPHH	4500
Db	4441	AAVRAGAQAARFTCTLSEAVPGEASWYINGAAVQDDSDMTVTADGSHQALLLRSQAOPHH	4500
Qy	4501	AGEVTFACRDAVASARLTVLGLPDDPEDAEVWASHSTVTLSWAAPMSDGGGGLCGYRVE	4560
Db	4501	AGEVTFACRDAVASARLTVLGLPDDPEDAEVWASHSTVTLSWAAPMSDGGGGLCGYRVE	4560
Qy	4561	VKEGATQWRLCHELVGPBECVVDGLAPGETYFRVAAGVPVGAGEPVHLPQTVRAEAPP	4620
Db	4561	VKEGATQWRLCHELVGPBECVVDGLAPGETYFRVAAGVPVGAGEPVHLPQTVRAEAPP	4620
Qy	4621	KPVPPQSPAPESRQVAAGEDVSLLEVVAAGEVIWHKGWERTQPCRFEVVSQGOQML	4680
Db	4621	KPVPPQSPAPESRQVAAGEDVSLLEVVAAGEVIWHKGWERTQPCRFEVVSQGOQML	4680
Qy	4681	VIKGFTAEDQGEYHCGLAQGSI CPAAATFOVALSPASVDEAPQPSLPPEAAQEGDLHLWL	4740
Db	4681	VIKGFTAEDQGEYHCGLAQGSI CPAAATFOVALSPASVDEAPQPSLPPEAAQEGDLHLWL	4740
Qy	4741	EALARKRMGREPTLDSISELPBEDGRSQRLOPEAEVAFPDLSEGYSTADELARTGDADL	4800
Db	4741	EALARKRMGREPTLDSISELPBEDGRSQRLOPEAEVAFPDLSEGYSTADELARTGDADL	4800
Qy	4801	SHTSSDDSESRAGTPSLTVYLKKAGRPCTSLASKVGAAPAPSVKPOQOQEPFLAARPPLG	4860
Db	4801	SHTSSDDSESRAGTPSLTVYLKKAGRPCTSLASKVGAAPAPSVKPOQOQEPFLAARPPLG	4860
Qy	4861	DLSTKOLGDPMSDKAAVKIOAAPKGYKVRKEMKQOQEPMSHTFGDTEAQVGDALRLECV	4920
Db	4861	DLSTKOLGDPMSDKAAVKIOAAPKGYKVRKEMKQOQEPMSHTFGDTEAQVGDALRLECV	4920
Qy	4921	VASKADVRAEWLKDGVELTDCRHHHIDOLGDGTCSLLIAGLDRADAGCYTCOVSNKFGQV	4980

4921	Db	V	A	K	A	D	V	R	A	R	W	L	K	D	G	V	E	L	T	D	G	R	H	H	I	D	L	Q	L	G	D	G	T	C	S	L	L	I	A	G	L	D	R	A	D	A	C	Y	T	C	V	S	N	K	F	Q	V	4	9	8
4981	Qy	T	H	A	C	V	V	V	S	G	S	E	A	S	S	G	E	L	D	A	P	R	R	A	R	L	H	R	L	F	R	T	K	S	P	A	E	V	S	D	E	L	F	L	S	A	D	E	G	P	A	5	0	4						
4981	Db	T	H	A	C	V	V	V	S	G	S	E	A	S	S	G	E	L	D	A	P	R	R	A	R	L	H	R	L	F	R	T	K	S	P	A	E	V	S	D	E	L	F	L	S	A	D	E	G	P	A	5	0	4						
5041	Qy	E	P	E	P	A	D	W	T	R	D	E	H	F	I	C	I	R	F	E	A	L	T	E	A	R	Q	A	V	T	R	F	Q	E	M	F	A	T	L	G	I	G	V	E	I	K	L	V	E	Q	P	R	V	E	M	5	1	0		
5041	Db	E	P	E	P	A	D	W	T	R	D	E	H	F	I	C	I	R	F	E	A	L	T	E	A	R	Q	A	V	T	R	F	Q	E	M	F	A	T	L	G	I	G	V	E	I	K	L	V	E	Q	P	R	V	E	M	5	1	0		
5101	Qy	C	I	S	K	E	T	P	A	P	V	V	P	E	P	L	S	L	T	S	A	A	P	V	L	T	E	L	Q	N	O	E	V	O	D	G	Y	P	V	S	F	D	C	V	T	G	O	M	P	S	V	R	P	5	1	6				
5101	Db	C	I	S	K	E	T	P	A	P	V	V	P	E	P	L	S	L	T	S	A	A	P	V	L	T	E	L	Q	N	O	E	V	O	D	G	Y	P	V	S	F	D	C	V	T	G	O	M	P	S	V	R	P	5	1	6				
5161	Qy	K	D	G	K	L	E	E	D	H	Y	M	I	N	E	D	Q	O	G	H	O	L	I	I	T	A	V	P	A	D	M	G	V	Y	R	C	L	A	E	N	S	M	G	S	V	S	T	K	A	E	L	R	V	D	L	T	S	2	2	0
5161	Db	K	D	G	K	L	E	E	D	H	Y	M	I	N	E	D	Q	O	G	H	O	L	I	I	T	A	V	P	A	D	M	G	V	Y	R	C	L	A	E	N	S	M	G	S	V	S	T	K	A	E	L	R	V	D	L	T	S	2	2	0
5221	Qy	T	D	Y	T	A	D	A	T	E	S	S	Y	F	S	A	Q	Y	L	S	R	O	E	G	T	E	S	T	D	E	G	O	L	P	O	V	V	E	L	R	D	L	O	V	A	P	C	T	R	L	A	K	F	5	2	8				
5221	Db	T	D	Y	T	A	D	A	T	E	S	S	Y	F	S	A	Q	Y	L	S	R	O	E	G	T	E	S	T	D	E	G	O	L	P	O	V	V	E	L	R	D	L	O	V	A	P	C	T	R	L	A	K	F	5	2	8				
5281	Qy	Q	L	K	V	G	Y	P	A	R	L	Y	W	F	K	D	G	Q	P	L	T	A	S	A	H	R	M	T	G	K	L	H	T	L	E	I	S	V	T	R	D	S	G	O	Y	A	A	I	S	N	A	M	G	5	3	4				
5281	Db	Q	L	K	V	G	Y	P	A	R	L	Y	W	F	K	D	G	Q	P	L	T	A	S	A	H	R	M	T	G	K	L	H	T	L	E	I	S	V	T	R	D	S	G	O	Y	A	A	I	S	N	A	M	G	5	3	4				
5341	Qy	A	A	Y	S	S	A	R	L	L	V	R	G	D	E	P	E	K	P	A	S	D																																						

QY 6061 ILIEDPDGSCALILDLSLTGVDGQYMCFAASAAGNCSTLTKILVQVPPRVFNKVRASPV 6120
DB 6061 ILIEDPDGSCALILDLSLTGVDGQYMCFAASAAGNCSTLTKILVQVPPRVFNKVRASPV 6120
QY 6121 EGEDAQFTTIEGAPYQIRWYKDGALLTGNKFQTLSEPRSGLLVIVIRAASKEDEGLY 6180
DB 6121 EGEDAQFTTIEGAPYQIRWYKDGALLTGNKFQTLSEPRSGLLVIVIRAASKEDEGLY 6180
QY 6181 ECENLNRLGSARASABLRISQPMLOAQEOCHREQLVAIVEDTTLR-----ADOEVT 6232
DB 6181 ECENLNRLGSARASABLRISQPMLOAQEOCHREQLVAIVEDTEQTKVPKTKVIIIEIT 6240
QY 6233 SVLXRLLGPKAPGPGTGDLTGPPCPRG-----APAL-----QETGSP- 6271
DB 6241 TVVKSPPRGQR-SPSKSPSPSPSRCSASPLRPGLLAPDLLYLPAGAGPRRPEAEPGQPV 6299
QY 6272 -PVTGTSAPAPVPRV-----PQPLLHEGP-----QPEAARAQEWTPVRMGA 6317
DB 6300 VPTLYVTEAEAHSPALPGLSGPQPKWVEETIEIVRVKMGPGQVSPTE--VPRSSSGH 6357
QY 6318 AW--PGAGTG 6325
DB 6358 LFTLPGATPG 6367
RESULT 4
US-10-077-130-2
; Sequence 2, Application US/10077130
; Publication No. US20020168742A1
; GENERAL INFORMATION:
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Acton, Susan L.
; TITLE OF INVENTION: 59079 and 12599, Protein Kinase Family
; FILE REFERENCE: MPI2001-047P(RCPL(M)
; CURRENT APPLICATION NUMBER: US/10/077,130
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269201
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2630
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-077-130-2
Query Match 33.2%; Score 13710; DB 13; Length 2630;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 5339 MGAAYSSARLLVRGDEPEEKPSADVHEQLVPPRMLEPFTPKVKKGSSITFSVKVEGRP 5398
DB 1 MGAAYSSARLLVRGDEPEEKPSADVHEQLVPPRMLEPFTPKVKKGSSITFSVKVEGRP 60
QY 5399 VPTVHMLREEAERGVLWTPGTPGYTVASSAQHSLVLLDVGRHQGTYYTCIASNAAGQA 5458
DB 61 VPTVHMLREEAERGVLWTPGTPGYTVASSAQHSLVLLDVGRHQGTYYTCIASNAAGQA 120
QY 5459 LCSASLHVSGLPKVEEQEKVEALISTFLOGTQTAISAGLETASFADLGGQKKEEPLAA 5518
DB 121 LCSASLHVSGLPKVEEQEKVEALISTFLOGTQTAISAGLETASFADLGGQKKEEPLAA 180
QY 5519 KEALGHLSLAEVGTETFEFLQKLTQITENVSAKITQAKLVQPGGSDSDSKTPSPASPRGR 5578
DB 181 KEALGHLSLAEVGTETFEFLQKLTQITENVSAKITQAKLVQPGGSDSDSKTPSPASPRGR 240
QY 5579 SRPSSIOESSSESDGDARGEIIFDIYVVTADYLPGLAEQDAITLREGQYVEVLDAAHPL 5638
DB 241 SRPSSIOESSSESDGDARGEIIFDIYVVTADYLPGLAEQDAITLREGQYVEVLDAAHPL 300
QY 5639 RNLVRTKTPSKSPRQGVSPAYLDRRLKLSPEWGAEEAFPEGAEVSEDEYKARLSSVI 5698

DB 301 RNLVRTKTPSKSPRQGVSPAYLDRRLKLSPEWGAEEAFPEGAEVSEDEYKARLSSVI 360
QY 5699 QELLSSEQAFVEELQFLOSHHLQHLERCPHPVPIAVAGOKAVI FRNVRDGRPHFSFLOEL 5758
DB 361 QELLSSEQAFVEELQFLOSHHLQHLERCPHPVPIAVAGOKAVI FRNVRDGRPHFSFLOEL 420
QY 5759 QOCOTDDDVANCFIKNOQAAPQYLEFLVGRVQAESVVVSTAIQBFYKKYAEALLAGDPS 5818
DB 421 QOCOTDDDVANCFIKNOQAAPQYLEFLVGRVQAESVVVSTAIQBFYKKYAEALLAGDPS 480
QY 5819 QPPPPPLQHYLEQPVVERVQYQALLKELIRNKARNRQNCALLEQAYAVVSALPORAENKL 5878
DB 481 QPPPPPLQHYLEQPVVERVQYQALLKELIRNKARNRQNCALLEQAYAVVSALPORAENKL 540
QY 5879 HVSILMENYPGTLEALGSPIRQGHFIWMEGAPGAPMPKGNHRVFLFRNHLVICKPRDS 5938
DB 541 HVSILMENYPGTLEALGSPIRQGHFIWMEGAPGAPMPKGNHRVFLFRNHLVICKPRDS 600
QY 5939 RTDTSVYVFRNMKLSISIDLNDQVEGDDRAPEVWQEREDSVRKYLLOARTAIKSSWVKE 5998
DB 601 RTDTSVYVFRNMKLSISIDLNDQVEGDDRAPEVWQEREDSVRKYLLOARTAIKSSWVKE 660
QY 5999 ICGIQORLALPVWRPDPFEBELADCTAELGETVKLACRVGTGPKPVI SWYKDGKAVQVDP 6058
DB 661 ICGIQORLALPVWRPDPFEBELADCTAELGETVKLACRVGTGPKPVI SWYKDGKAVQVDP 720
QY 6059 HHILIEDPDGSCALILDLSLTGVDGQYMCFAASAAGNCSTLTKILVQVPPRVFNKVRASPV 6118
DB 721 HHILIEDPDGSCALILDLSLTGVDGQYMCFAASAAGNCSTLTKILVQVPPRVFNKVRASPV 780
QY 6119 FVEGEDAQFTTIEGAPYQIRWYKDGALLTGNKFQTLSEPRSGLLVIVIRAASKEDEGL 6178
DB 781 FVEGEDAQFTTIEGAPYQIRWYKDGALLTGNKFQTLSEPRSGLLVIVIRAASKEDEGL 840
QY 6179 LYECELVNRLGSARASABLRISQPMLOAQEOCHREQLVAAVEDTTLRADQEVTSVLKRL 6238
DB 841 LYECELVNRLGSARASABLRISQPMLOAQEOCHREQLVAAVEDTTLRADQEVTSVLKRL 900
QY 6239 LGPKAPGSTGDLTGPGCPRGAPALQETGSPVVTGTSBPAPVPPRPVPPOLLHEGPEQE 6298
DB 901 LGPKAPGSTGDLTGPGCPRGAPALQETGSPVVTGTSBPAPVPPRPVPPOLLHEGPEQE 960
QY 6299 PEAIARAQEWTPVRMEEGAAMPAGAGTCELLVDVSHVHVRETTQRTYTYQAITHTARPPS 6358
DB 961 PEAIARAQEWTPVRMEEGAAMPAGAGTCELLVDVSHVHVRETTQRTYTYQAITHTARPPS 1020
QY 6359 MOVTIEDVQAGTGTGTAQFEAIIEGDPQPSVTWYKDSVQLVDSTRLSQOQEGTYSVLVRH 6418
DB 1021 MOVTIEDVQAGTGTGTAQFEAIIEGDPQPSVTWYKDSVQLVDSTRLSQOQEGTYSVLVRH 1080
QY 6419 VASKDAGVYTCLAQNTGGVLCKAELLVLGDDNEPDESEKSHRRKLSHFYEVKEEIGRGV 6478
DB 1081 VASKDAGVYTCLAQNTGGVLCKAELLVLGDDNEPDESEKSHRRKLSHFYEVKEEIGRGV 1140
QY 6479 FGFVKRVQHKGNKILCAAKFIPLRSRTRAQAYRERDILAAALSHPLVTGLDQFETKTLI 6538
DB 1141 FGFVKRVQHKGNKILCAAKFIPLRSRTRAQAYRERDILAAALSHPLVTGLDQFETKTLI 1200
QY 6539 LILELCSSEBLLDRLYRKGVVTEAEVKVYIQQLVEGLHYLHSHGVHLHDKPSNLMVHP 6598
DB 1201 LILELCSSEBLLDRLYRKGVVTEAEVKVYIQQLVEGLHYLHSHGVHLHDKPSNLMVHP 1260
QY 6599 AREDIKICDPGFAQNTTPAELQPSQSGPFSVPEIITQONPVSEASDIWANGVISYLSLT 6658
DB 1261 AREDIKICDPGFAQNTTPAELQPSQSGPFSVPEIITQONPVSEASDIWANGVISYLSLT 1320
QY 6659 CSSPFPAGESDRATLLNVLEGRVSNWSPMAHLSEDADKDFIKATLQAPQAPRPSAAQCLSH 6718
DB 1321 CSSPFPAGESDRATLLNVLEGRVSNWSPMAHLSEDADKDFIKATLQAPQAPRPSAAQCLSH 1380
QY 6719 PWFLLKSNPAAEAFINTKOLKFLLRARWORSWMSYKSLVMSRIPPELLRGPDPSPSLGV 6778

QY	5913	MPWKGNHRHVF	FRNHLVI	CKPRDRSRD	TDTSYV	FRNNMKLSS	IDLNDQ	VEGDDRAFEVW	5972
Db	541	MPWKGNHRHVF	FRNHLVI	CKPRDRSRD	TDTSYV	FRNNMKLSS	IDLNDQ	VEGDDRAFEVW	600
QY	5973	QEREDSVRKYLL	LOARTALII	KSSWKEICGI	QORLAL	PVMRPDP	FEELH	ADCTAELGTVK	6032
Db	601	QEREDSVRKYLL	LOARTALII	KSSWKEICGI	QORLAL	PVMRPDP	FEELH	ADCTAELGTVK	660
QY	6033	LACRVGTGPKPVI	SWYKDGKAVQ	DPHHILIED	PDGSCAL	ILDSLT	IGVDSQYMCFAASA	6092	
Db	661	LACRVGTGPKPVI	SWYKDGKAVQ	DPHHILIED	PDGSCAL	ILDSLT	IGVDSQYMCFAASA	720	
QY	6093	AGNCSTLGKILV	QVPPRVFNKVR	ASPFVEGEDA	QFTCTIEGAPY	QOIRWKDGALLTGN	6152		
Db	721	AGNCSTLGKILV	QVPPRVFNKVR	ASPFVEGEDA	QFTCTIEGAPY	QOIRWKDGALLTGN	780		
QY	6153	KFOTLSBPRSGLLV	IVIRAAASKED	GLGYCECLVNRLG	SARASAEIR	IOSPMLQAOEOCHR	6212		
Db	781	KFOTLSBPRSGLLV	IVIRAAASKED	GLGYCECLVNRLG	SARASAEIR	IOSPMLQAOEOCHR	840		
QY	6213	EOLVAAVEDTTLER	ADQEVTSVLKRL	GLPKAPG	STGDLTGPG	FCPRGAPALQETGSOQP	6272		
Db	841	EOLVAAVEDTTLER	ADQEVTSVLKRL	GLPKAPG	STGDLTGPG	FCPRGAPALQETGSOQP	900		
QY	6273	VTGTSEAPAVP	PRVPOPLLHG	PQBPBEA	IARAQEW	TPVIMGEAANWPGAGTGELLMDVH	6332		
Db	901	VTGTSEAPAVP	PRVPOPLLHG	PQBPBEA	IARAQEW	TPVIMGEAANWPGAGTGELLMDVH	960		
QY	6333	SHVRETTRQYTY	QAIIDHTHARP	SPSQVTTIED	VQAOTGGT	TAQFEAII	IEGDPQSPVTTYK	6392	
Db	961	SHVRETTRQYTY	QAIIDHTHARP	SPSQVTTIED	VQAOTGGT	TAQFEAII	IEGDPQSPVTTYK	1020	
QY	6393	DSVQLVDS	TRLSQQOEG	TTYSLVLR	HVASKDAG	YVYTCLAQNTG	QGVLCKAELLVLGGDNE	6452	
Db	1021	DSVQLVDS	TRLSQQOEG	TTYSLVLR	HVASKDAG	YVYTCLAQNTG	QGVLCKAELLVLGGDNE	1080	
QY	6453	PDSEKQSHRRK	LKHSFYEVKE	BIEGRGV	GFVKRVQ	HGKNKILCAAK	FPLBSRTRAQAYRE	6512	
Db	1081	PDSEKQSHRRK	LKHSFYEVKE	BIEGRGV	GFVKRVQ	HGKNKILCAAK	FPLBSRTRAQAYRE	1140	
QY	6513	RDILAAALSHPL	VTGLLOQ	FOFETRK	TLII	LILCSSEELL	DLRYKGVVTEAEVKVYIQOLV	6572	
Db	1141	RDILAAALSHPL	VTGLLOQ	FOFETRK	TLII	LILCSSEELL	DLRYKGVVTEAEVKVYIQOLV	1200	
QY	6573	EGLHYLHSHG	VHLDDI	KPSNILMV	HPAREDI	KICDGF	FAQNTIPAEILOF	QSYGSPFVSP	6632
Db	1201	EGLHYLHSHG	VHLDDI	KPSNILMV	HPAREDI	KICDGF	FAQNTIPAEILOF	QSYGSPFVSP	1260
QY	6633	EIIQOONPV	SEASDIWANG	VTISYLSL	TCSPPAGES	DRATILNV	LEGVSNSSPMAHLSE	6692	
Db	1261	EIIQOONPV	SEASDIWANG	VTISYLSL	TCSPPAGES	DRATILNV	LEGVSNSSPMAHLSE	1320	
QY	6693	DAKDFIKATL	QARAQAP	SAQCLSH	SPWFLKSM	PAEAEHINT	KOLKFLLLARSRWORS	6752	
Db	1321	DAKDFIKATL	QARAQAP	SAQCLSH	SPWFLKSM	PAEAEHINT	KOLKFLLLARSRWORS	1380	
QY	6753	SYKSIILVMRSI	PEILLRGP	PPSPSLGV	VARHLC	RDTGGSSSSSS	SSSDNELAPFARAKSLPPS	6812	
Db	1381	SYKSIILVMRSI	PEILLRGP	PPSPSLGV	VARHLC	RDTGGSSSSSS	SSSDNELAPFARAKSLPPS	1440	
QY	6813	PVTHSPILLH	PRGFLRPS	ASLPEEAE	ASERSTE	APAPAS	PEGAGPPAAOC	CVPRHSVIRS	6872
Db	1441	PVTHSPILLH	PRGFLRPS	ASLPEEAE	ASERSTE	APAPAS	PEGAGPPAAOC	CVPRHSVIRS	1500
QY	6873	LFYHQAGES	PEHGALAF	SGSRRH	PARRHLL	KGGYIAGAL	PGLREPLMEHR	VLEEEAAREE	6932
Db	1501	LFYHQAGES	PEHGALAF	SGSRRH	PARRHLL	KGGYIAGAL	PGLREPLMEHR	VLEEEAAREE	1560
QY	6933	QATLLAKAP	FETALRULP	PASG	THLAPGH	SHSLHSD	SPSTPRPS	SEACGEAQRULPSAPSGG	6992
Db	1561	QATLLAKAP	FETALRULP	PASG	THLAPGH	SHSLHSD	SPSTPRPS	SEACGEAQRULPSAPSGG	1620

Qy	6993	APIRDMGHPQGSKOLPSTGGHPTGTAQPERSPDPSPMGQAPAPCPHPXOGSAPQGCSCPHPA	7052
Db	1621	APIRDMGHPQGSKOLPSTGGHPTGTAQPERSPDPSPMGQAPAPCPHPXOGSAPQGCSCPHPA	1680
Qy	7053	VAPCPPGSPFGSGCKEAPLVPSPFGLQOPAPAPAKASPPDLSKMGCPGDISLPGRPKPG	7112
Db	1681	VAPCPPGSPFGSGCKEAPLVPSPFGLQOPAPAPAKASPPDLSKMGCPGDISLPGRPKPG	1740
Qy	7113	PCSSPGSQAASSSQVSSLRVGSSQVTEPGPSLDAEGWTQEAEDLSDSPTTLQRQEOQA	7172
Db	1741	PCSSPGSQAASSSQVSSLRVGSSQVTEPGPSLDAEGWTQEAEDLSDSPTTLQRQEOQA	1800
Qy	7173	TMRXFSLGGRGYAGVAGYGTAFGGDAGMLCQGPMMWARIAWAVSQSEBEEQEEARAES	7232
Db	1801	TMRXFSLGGRGYAGVAGYGTAFGGDAGMLCQGPMMWARIAWAVSQSEBEEQEEARAES	1860
Qy	7233	QSEBQOEARAESPLUPQVARSAPVEVGRAPTRSSPEPTMEDIGQVSLVQIRDLSGDAEAA	7292
Db	1861	QSEBQOEARAESPLUPQVARSAPVEVGRAPTRSSPEPTMEDIGQVSLVQIRDLSGDAEAA	1920
Qy	7293	DTISLDSIVDPAYLNLSDLYDIKYLPFFPMIRKVPKSAQPEPPSPMAEELAEPEPT	7352
Db	1921	DTISLDSIVDPAYLNLSDLYDIKYLPFFPMIRKVPKSAQPEPPSPMAEELAEPEPT	1980
Qy	7353	WPWPEGLPHAGLIEETESDDVALLAAEAVGRKRWKSSPSRSLPHFPGRHLPDPEAEL	7412
Db	1981	WPWPEGLPHAGLIEETESDDVALLAAEAVGRKRWKSSPSRSLPHFPGRHLPDPEAEL	2040
Qy	7413	GLRRVKVASVEHISRILKGRPEGLEKEGPRKPGLASFRLSGLKSWDRAPFRLRLSDE	7472
Db	2041	GLRRVKVASVEHISRILKGRPEGLEKEGPRKPGLASFRLSGLKSWDRAPFRLRLSDE	2100
Qy	7473	TVVLGQSVTILACQVSAQPAQAATWSKDGAPLEBSSRVLISATLKNFQLLTILVVAEDLIG	7532
Db	2101	TVVLGQSVTILACQVSAQPAQAATWSKDGAPLEBSSRVLISATLKNFQLLTILVVAEDLIG	2160
Qy	7533	VYTCVSNALGTVTTTGLVKRAERPSSSPCPDIGEVYADGVLLVWKPVSYPGVTVIVOC	7592
Db	2161	VYTCVSNALGTVTTTGLVKRAERPSSSPCPDIGEVYADGVLLVWKPVSYPGVTVIVOC	2220
Qy	7593	SLEGGSWTTLASDIFDCCYLTSKLSRGTVTFRACVSKAGMPYSSPSEQVLLGPGSHL	7652
Db	2221	SLEGGSWTTLASDIFDCCYLTSKLSRGTVTFRACVSKAGMPYSSPSEQVLLGPGSHL	2280
Qy	7653	ASEBESQGRSAQPLPSTKTFAFOTQORGFSVVROCWEKASGRALAAKIIIPYHPKDKTA	7712
Db	2281	ASEBESQGRSAQPLPSTKTFAFOTQORGFSVVROCWEKASGRALAAKIIIPYHPKDKTA	2340
Qy	7713	VLREYALKGRPHPLAQLHAAVYLSPRHLVLIILELCSGPPELLPCLAEARASYSESEVKDYL	7772
Db	2341	VLREYALKGRPHPLAQLHAAVYLSPRHLVLIILELCSGPPELLPCLAEARASYSESEVKDYL	2400
Qy	7773	WQMLSAQYLIHQHILHLDLRSNNMIITEYNLLKVVDLGNQASLSQEKVLPSPDKFDXYLE	7832
Db	2401	WQMLSAQYLIHQHILHLDLRSNNMIITEYNLLKVVDLGNQASLSQEKVLPSPDKFDXYLE	2460
Qy	7833	TMAPELLGCGAVPOTDIIWAGYTAIMLSABEYPUSSEGGARDLQRLKGLVRLSRCYAG	7892
Db	2461	TMAPELLGCGAVPOTDIIWAGYTAIMLSABEYPUSSEGGARDLQRLKGLVRLSRCYAG	2520
Qy	7893	LSGGAVAFRLSTLCAQPWGRPCASSCLQCPWLITEGPACSRPAPVTFPTARLRFVVRNRE	7952
Db	2521	LSGGAVAFRLSTLCAQPWGRPCASSCLQCPWLITEGPACSRPAPVTFPTARLRFVVRNRE	2580
Qy	7953	KRRALLYKRHNLAQVR	7968
Db	2581	KRRALLYKRHNLAQVR	2596

RESULT 6
US-10-093-463-72
; Sequence 72, Application US/10093463
; Publication No. US20030208039A1

GENERAL INFORMATION:
APPLICANT: Padigaru, Muralidhara
APPLICANT: Shenoy, Suresh
APPLICANT: Kekuda, Ramesh
APPLICANT: Gusev, Vladimir
APPLICANT: Pochart, Pascal
APPLICANT: Zhong, Mei
APPLICANT: Rastelli, Luca
APPLICANT: Mezes, Peter
APPLICANT: Smithson, Glennnda
APPLICANT: Guo, Xiaojia
APPLICANT: Gerlach, Valerie
APPLICANT: Casman, Stacie
APPLICANT: Boldog, Ferenc
APPLICANT: Li, Li
APPLICANT: Zerhusen, Bryan
APPLICANT: Tchernev, Velizar
APPLICANT: Gangolli, Esha
APPLICANT: Vernet, Corine
APPLICANT: Pena, Carol
APPLICANT: Burgess, Catherine
APPLICANT: Liu, Xiaohong
APPLICANT: Spytek, Kimberly
APPLICANT: Gorman, Linda
APPLICANT: Spaderna, Steven
APPLICANT: Voss, Edward
APPLICANT: Malyankar, Uriel
APPLICANT: Anderson, David
APPLICANT: Patturajan, Meera
APPLICANT: Miller, Charles
APPLICANT: Taupier, Raymond J. Jr.
TITLE OF INVENTION: No. US20030208039A1el Antibodies that Bind to Antigenic Polypeptide
TITLE OF INVENTION: Encoding The Antigen
FILE REFERENCE: 21402-290A (Cura 590At)
CURRENT APPLICATION NUMBER: US/10/093,463
CURRENT FILING DATE: 2002-06-24
PRIOR APPLICATION NUMBER: 60/283,675
PRIOR FILING DATE: 2001-04-14
PRIOR APPLICATION NUMBER: 60/338,092
PRIOR FILING DATE: 2001-12-03
PRIOR APPLICATION NUMBER: 60/274,281
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,101
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/325,681
PRIOR FILING DATE: 2001-09-27
PRIOR APPLICATION NUMBER: 60/304,354
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/279,995
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 60/294,899
PRIOR FILING DATE: 2001-05-31
PRIOR APPLICATION NUMBER: 60/287,424
PRIOR FILING DATE: 2001-04-30
PRIOR APPLICATION NUMBER: 60/299,027
PRIOR FILING DATE: 2001-06-18
PRIOR APPLICATION NUMBER: 60/309,198
PRIOR FILING DATE: 2001-07-31
PRIOR APPLICATION NUMBER: 60/281,194
PRIOR FILING DATE: 2001-04-04
PRIOR APPLICATION NUMBER: 60/274,194
PRIOR FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 60/274,849
PRIOR FILING DATE: 2001-03-09
PRIOR APPLICATION NUMBER: 60/330,380
PRIOR FILING DATE: 2001-10-18
PRIOR APPLICATION NUMBER: 60/275,235
PRIOR FILING DATE: 2001-03-12
PRIOR APPLICATION NUMBER: 60/288,342
PRIOR FILING DATE: 2001-05-03
PRIOR APPLICATION NUMBER: 60/275,578
PRIOR FILING DATE: 2001-03-13
NUMBER OF SEQ ID NOS: 370

SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 72
LENGTH: 4691
TYPE: PRT
ORGANISM: Homo sapiens
US-10-093-463-72

Query Match 25.5%; Score 10519.5; DB 15; Length 4691;
Best Local Similarity 51.1%; Pred. No. 0;
Matches 2367; Conservative 118; Mismatches 342; Indels 1807; Gaps 66;
QY 2595 MPLYNDSPHEISHKGRRTTLVLSIQRADAGIVRASSLKVSTASRLVVRVKPVVFLKALD 2654
Db 1 MPLYNDSPHEISHKGRRTTLVLSIQRADAGIVRASSLKVSTASRLVVRVKPVVFLKALD 60
QY 2655 DLSAEERGTALQCEVSDPEAHVVRKDGVLGPSDKYDFLHTACTRGLVHVDVSPEDAG 2714
Db 61 DLSAEERGTALQCEVSDPEAHVVRKDGVLGPSDKYDFLHTACTRGLVHVDVSPEDAG 120
QY 2715 LYTHVGSEETRARVRVHDLHVGITKRLKTMVELEGESCSPECVLSHESASDPAMWTGG 2774
Db 121 LYTHVGSEETRARVRVHDLHVGITKRLKTMVELEGESCSPECVLSHESASDPAMWTGG 180
QY 2775 KTVGSSSRFQATROGRKYILVVRGAAPSDAGEWVFSVRLTSKSLIVRERPAAIKPLE 2834
Db 181 KTVGSSSRFQATROGRKYILVVRGAAPSDAGEWVFSVRLTSKSLIVRERPAAIKPLE 240
QY 2835 DQWVAPGEDVELRCELSRAGTPVHMLKDKKAIKRSQKDVVCEGTMAMLVIRGASLKDAG 2894
Db 241 DQWVAPGEDVELRCELSRAGTPVHMLKDKKAIKRSQKDVVCEGTMAMLVIRGASLKDAG 300
QY 2895 EYTCVEASKSTASLHVBEKANCFTTELTNLQVEBKGTAVFTCKTEHPAATVTRKGLLE 2954
Db 301 EYTCVEASKSTASLHVBEKANCFTTELTNLQVEBKGTAVFTCKTEHPAATVTRKGLLE 360
QY 2955 LRASGKHQPSQEGTLRLTISALEKADSDTYTCDIGQAQSRQAQLVQGRVRIIEDLEV 3014
Db 361 LRASGKHQPSQEGTLRLTISALEKADSDTYTCDIGQAQSRQAQLVQGRVRIIEDLEV 420
QY 3015 DVQEGSSATFRCRISPANYEPVHWFDPKTPHLANELNIDAPGGYHVLTLRQALKDSG 3074
Db 421 DVQEGSSATFRCRISPANYEPVHWFDPKTPHLANELNIDAPGGYHVLTLRQALKDSG 480
QY 3075 TIYFAGQORASALRVTEKPSVFSRELTDTATITEGEDTLVCETSTCDIPMCWTGDGKT 3134
Db 481 TIYFAGQORASALRVTEKPSVFSRELTDTATITEGEDTLVCETSTCDIPMCWTGDGKT 540
QY 3135 LRGSARCOLSHEGHAQLITGATLQDSGRYKCEAGGACSSIVRVHARPVRFOALKDOL 3194
Db 541 LRGSARCOLSHEGHAQLITGATLQDSGRYKCEAGGACSSIVRVHARPVRFOALKDOL 600
QY 3195 EVLEGGAATLRVLSVAAPVKWCYGNVLRPGDKYSLRQEGAMLELVVRNLRPQDSGRY 3254
Db 601 EVLEGGAATLRVLSVAAPVKWCYGNVLRPGDKYSLRQEGAMLELVVRNLRPQDSGRY 660
QY 3255 SCSFGDQTTSATLTVTALPAQFIGKLRNKEATEGATATLRCELSKAAVPVWRKGSSETLAD 3314
Db 661 SCSFGDQTTSATLTVTALPAQFIGKLRNKEATEGATATLRCELSKAAVPVWRKGSSETLAD 720
QY 3315 GDRYCLRODGAMCELOIRGLAMVDAEYSCVCGEERTSASLTIRPMPAHFIRLHRQESI 3374
Db 721 GDRYCLRODGAMCELOIRGLAMVDAEYSCVCGEERTSASLTIRPMPAHFIRLHRQESI 780
QY 3375 EGATATLRCELSKAAVPVWRKGSSETLADLRQDGAVALCELOICGLAVADAGEYSCVC 3434
Db 781 EGATATLRCELSKAAVPVWRKGSSETLADLRQDGAVALCELOICGLAVADAGEYSCVC 840
QY 3435 GEERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLCBELSKVAPVWRKGPENLRDGRY 3494
Db 841 GEERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLCBELSKVAPVWRKGPENLRDGRY 900
QY 3495 ILRQEGTRCELOICGLAMADAGEYLCVCGEERTSATLIRALPASFIEDVKNQAREGAT 3554

Db	901	ILRQEGTRCELQICGLAMADAGEYLCVCGQERTSATLTITRALPARFIEDVKNQAREGAT	960
Qy	3555	AVLQCELSAAPVEMWKGSETLRDGRYSLRQDGTKECELQIRGLAMADTGEYSCVCGQER	3614
Db	961	AVLQCELSAAPVEMWKGSETLRDGRYSLRQDGTKECELQIRGLAMADTGEYSCVCGQER	1020
Qy	3615	TSAMLTIVRALPIKFTBGLRNEBATEGATAVLQCELSKAPVEMWKGHETLRDGRHSLRQ	3674
Db	1021	TSAMLTIVRALPIKFTBGLRNEBATEGATAVLQCELSKAPVEMWKGHETLRDGRHSLRQ	1080
Qy	3675	DGARCELOJRGVAEDDAGYLCWCGKERTSAMLTVRAMPSEKIEGLRNEEATEGDTATLW	3734
Db	1081	DGARCELOJRGVAEDDAGYLCWCGKERTSAMLTVRAMPSEKIEGLRNEEATEGDTATLW	1140
Qy	3735	CELSKAAAPVEMWKGHETLRDGRHSLRQDGRSCELCQIRGLAVVDAGEYSCVCGQERTSAT	3794
Db	1141	CELSKAAAPVEMWKGHETLRDGRHSLRQDGRSCELCQIRGLAVVDAGEYSCVCGQERTSAT	1200
Qy	3795	LTVRALPARFIEDVKNQAREGATAVLQCELSKAAAPVEMWKGSETLRDGRYSLRQDGTGR	3854
Db	1201	LTVRALPARFIEDVKNQAREGATAVLQCELSKAAAPVEMWKGSETLRDGRYSLRQDGTGR	1260
Qy	3855	CELOIHGLSVADTGEYSCVCGQERTSATLTIVR-----	3886
Db	1261	CELOIHGLSVADTGEYSCVCGQERTSATLTIVRALPARFTQDLKTKASEGATATLOCELS	1320
Qy	3887	-----	3886
Db	1321	KVAPVEMWKGSETLRDGRYSLRQDGTGRCELQIHDLVSADAGEYSCMCGQERTSATLTIVR	1380
Qy	3887	-----	3886
Db	1381	ALPARFTEGLRNEEAMEGATATLOCELSKAAAPVEMWKGLEALRDGDKYSLRDGAVCELCQ	1440
Qy	3887	-----	3886
Db	1441	IHGLAMADNVVYSCVCGQERTSATLTIVRALPARFIEDMRNOKATEGATVTLQCKLRKAAP	1500
Qy	3887	-----	3886
Db	1501	VEWKGPNLTLDGRYSLRQDGTSCCELQIRGLVIADAGEYSCIQERTSATLTIVRALPA	1560
Qy	3887	-----	3886
Db	1561	RFIEDVRNHEATEGATAVLQCELSKAAAPVEMWKGSETLRDGRYSLRQDGTGRCELQIRGL	1620
Qy	3887	-----	3886
Db	1621	AVEDTGEYLCVCGQERTSATLTIVRALPARFIDNMTNQAREGATATLHCELSKVAPVEMW	1680
Qy	3887	-----	3886
Db	1681	KGPETLRDGRHSLRQDGRSCELCQIRGLAVVDAGEYSCVCGQERTSATLTIVRALPARFIE	1740
Qy	3887	-----	3886
Db	1741	DVKNQAREGATAVLQCELSKAAAPVEMWKGSETLRDGRYSLRQDGTGRCELQIHGLSVAD	1800
Qy	3887	-----	3886
Db	1801	TGEYSCVCGQERTSATLTIVRALPARFTQDLKTKASEGATATLOCELSKVAPVEMWKGPE	1860
Qy	3887	-----	3886
Db	1861	TLRDGRYSLRQDGTGRCELQIHDLVSADAGEYSCMCGQERTSATLTIVRDCHTLVMPHY	1920
Qy	3887	-----	3886
Db	1921	FQLPGLLKEPETHIYIQIPSPVILFTBGLRNEEAMEGATATLOCELSKAAAPVEMWKGLE	1980
Qy	3887	-----	3886
Db	1981	ALRDGDKYSLRDGAVCELCQIHGLAMADNVVYSSILPARFIEDMRNOKATEGATVTLQCKL	2040
Qy	3887	-----	3886
Db	2041	RKAAPVEMWKGPNLTLDGRYSLRQDGTSCCELQIRGLVIADAGEYSCICEQERTSATLTIV	2100
Qy	3887	-----	3886
Db	2101	RALPARFIEDVRNHEATEGATAVLQCELSKAAAPVEMWKGSETLRDGRYSLRQDGTGRCEL	2160
Qy	3887	-----	3886
Db	2161	QIRGLAVEDTGEYLCVCGQERTSATLTIVRALPARFIDNMTNQAREGATATLHCELSKVA	2220
Qy	3887	-----	3886
Db	2221	PVEMWKGSETLRDGRHSLRQDGTGRCELQIRGLSVADAGEYSCVCGQERTSATLTIRALP	2280
Qy	3887	-----	3886
Db	2281	AKFTKGLRNEBATEGATAMLOCELSKVAPVEMWKGSETLRDGRYNLRQDGTGRCELQIHG	2340
Qy	3887	-----	3886
Db	2341	LSVADTGEYSCVCGQERTSATLTIVKAPQVPREPLQSLQAEBSGSTATLQCELSBPTATV	2400
Qy	3921	WSKGLQLOQANGREPRELQCGCTABELVLQDLQREDTGEYTCCTCGSOATSATLTIVTAAPVRF	3981
Db	2401	WSKGLQLOQANGREPRELQCGCTABELVLQDLQREDTGEYTCCTCGSOATSATLTIVTAAPVRF	2460
Qy	3981	LRELQHOEVDGEGTAHLCCCELSRAGASVEMWKGSLQLPCCAKYQMVODGAAAEALLVRGVE	4041
Db	2461	LRELQHOEVDGEGTAHLCCCELSRAGASVEMWKGSLQLPCCAKYQMVODGAAAEALLVRGVE	2520
Qy	4041	QEDAGDVTCTDGTGTSNASLSVRPRPKFTRLOSLEQETGDIARLCCQLSDASGAVVQ	4101
Db	2521	QEDAGDVTCTDGTGTSNASLSVRGGR-----	2549
Qy	4101	WLKGEVLEHAGPKYEMRSQATRELLIHOLEAKDTGEYACVTCGQKTAASLRVTEPEVTI	4161
Db	2550	-----	2600
Qy	4161	VRGLVDAEVTDAEDVEFSCEVSRAGATGVQWQCLQGLFLOSNEVTEVAVRDGRHITLRKG	4221
Db	2601	VRGLVDAEVTDAEDVEFSCEVSRAGATGVQWQCLQGLFLOSNEVTEVAVRDGRHITLRKG	2660
Qy	4221	VTPEADGTVSGFHLGNHASSAQLTVRAPEVITLPLQDVQVQSEGDASFQCRSLSRASGQEA	4281
Db	2661	VTPEADGTVSGFHLGNHASSAQLTVRAPEVITLPLQDVQVQSEGDASFQCRSLSRASGQEA	2701
Qy	4281	RWALGGVPLQANEMNDITVEQGTLLHLLTLHKVTLLEDAGTVSFHVGTCSSEALQKVT	4337
Db	2701	-----	2756
Qy	4337	-----	4337
Db	2757	CLVRGLQNVDFAGEVATFSCEDGQPSAIAVRDGIHFHSLMLSGLVADSGTVIFRAGPLV	2816
Qy	4338	-----	4337
Db	2817	STAKLLIKDPVVEVVSAMQDLAVEEGGSABELLCOYSRPVQATWQKDBREVHTDGRHVIE	2876
Qy	4338	-----	4360
Db	2877	QDMNVARLTFRPALPCDSGIYSCAAAGTRVVALLVQVAKNTVVRGLENVEALEGEALFE	2936
Qy	4361	COLSQOPEVAATHLLDDEPVRTSENABEVFFENGRLHLLLLKNLRPODSKRVTFLAGDMV	4420
Db	2937	COLSQOPEVAATHLLDDEPVRTSENABEVFFENGRLHLLLLKNLRPODSKRVTFLAGDMV	2996
Qy	4421	TSAPLTVR-----	4472
Db	2997	TSAPLTVRGCAVLVQWRLELLEPLKNAAVRAGAQRFTCTLSEAVPVGEASWYINGAA	3056

[illegible]

D5	4002	-----YAPSATETQS-----PAP---WSPRAPE---	4021
QY	5414	LWIGPDTPGYTVASSAQOHSVLVDDVCRHQGYYTCTIASNAQAALCSASLHVSGLPKVE	5473
D5	4022	-----PEHPG--VPSLAFPSA-----RACAAAGPYGPSRAA	4051
QY	5474	EQEKVKCALISTFLOQTTOAISAGLETFADLGQORKEEPLAAKEALGHLSLAEVGT	5533
D5	4052	EAARRRPADSTAFLP-SVRMAA-----PDLSTNLQEEATCA-----	4087
QY	5534	EFLOKLTQSITEMVMSAKITQAKLVPGGDSEDESKTPSAS--PRHGSRSPSSI----	5585
D5	4088	ICLDYFTDPVMTDCGHNFCECIRRCWGOWPEATRAPSAAASCRRGTCTARLLRWPRW	4147
QY	5586	QESSSESSEDGARGETFDIIVVTADVLPGLCAE-QDAITL-----REGQYVEVDAAHPL	5638
D5	4148	RGACTRRRRRRR-----VPAHREPLAAF CGDELRLLLCAACERSGBH-----	4189
QY	5639	RWL-----VRTKPTKSSPRSQG-----WVSPAYLDRLRK-LSPB-----	5671
D5	4190	-WAHVGRCTRTRPKTSRPLEAGTWAAANETLLSGAKLEKSLEHLURKOMQDALLFOAQDET	4248
QY	5672	---WGAAAEAPFFEGEAVSEDEYKARLSVSVTIQLLSSEQAPVEELOFLQSHHL----QHLE	5724
D5	4249	CVLWQABEDGGEQQNVLREFE---LRURLLAEGTTAAAEAGEELKQSAHLAELTAELE	4305
QY	5725	RCPHVPJAVAQKA	5738
D5	4306	R--PLPAACAGAAA	4317

RESULT 7

```

RESULT 7
; Sequence 74, Application US/10093463
; Publication No. US20020208039A1
; GENERAL INFORMATION:
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shenoy, Suresh
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Gusev, Vladimir
; APPLICANT: Pochart, Pascal
; APPLICANT: Zhong, Mei
; APPLICANT: Rastelli, Luca
; APPLICANT: Mezes, Peter
; APPLICANT: Smithson, Glennda
; APPLICANT: Guo, Xiaojia
; APPLICANT: Gerlach, Valerie
; APPLICANT: Casman, Stacie
; APPLICANT: Boldog, Ferenc
; APPLICANT: Li, Li
; APPLICANT: Zernusen, Bryan
; APPLICANT: Tchernev, Velizar
; APPLICANT: Gangolli, Esha
; APPLICANT: Vernet, Corinne
; APPLICANT: Pena, Carol
; APPLICANT: Burgess, Catherine
; APPLICANT: Liu, Xiaohong
; APPLICANT: Spytek, Kimberly
; APPLICANT: Gorman, Linda
; APPLICANT: Spaderna, Steven
; APPLICANT: Voss, Edward
; APPLICANT: Malyankar, Uriel
; APPLICANT: Anderson, David
; APPLICANT: Patturajan, Meera
; APPLICANT: Miller, Charles
; APPLICANT: Taupier, Raymond J. Jr.
; TITLE OF INVENTION: No. US200302080
; FILE OF INVENTION: Encoding The A
; FILE REFERENCE: 21402-290A (Cura 59
; CURRENT APPLICATION NUMBER: US/10/
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 60/283.67

```



```
/ PRIOR FILING DATE: 2001-04-14
/ PRIOR APPLICATION NUMBER: 60/338,092
/ PRIOR FILING DATE: 2001-12-03
/ PRIOR APPLICATION NUMBER: 60/274,281
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 60/274,101
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 60/325,681
/ PRIOR FILING DATE: 2001-09-27
/ PRIOR APPLICATION NUMBER: 60/304,354
/ PRIOR FILING DATE: 2001-07-10
/ PRIOR APPLICATION NUMBER: 60/279,995
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 60/294,899
/ PRIOR FILING DATE: 2001-05-31
/ PRIOR APPLICATION NUMBER: 60/287,424
/ PRIOR FILING DATE: 2001-04-30
/ PRIOR APPLICATION NUMBER: 60/299,027
/ PRIOR FILING DATE: 2001-06-18
/ PRIOR APPLICATION NUMBER: 60/309,198
/ PRIOR FILING DATE: 2001-07-31
/ PRIOR APPLICATION NUMBER: 60/281,194
/ PRIOR FILING DATE: 2001-04-04
/ PRIOR APPLICATION NUMBER: 60/274,194
/ PRIOR FILING DATE: 2001-03-08
/ PRIOR APPLICATION NUMBER: 60/274,849
/ PRIOR FILING DATE: 2001-03-09
/ PRIOR APPLICATION NUMBER: 60/330,380
/ PRIOR FILING DATE: 2001-10-18
/ PRIOR APPLICATION NUMBER: 60/275,235
/ PRIOR FILING DATE: 2001-03-12
/ PRIOR APPLICATION NUMBER: 60/288,342
/ PRIOR FILING DATE: 2001-05-03
/ PRIOR APPLICATION NUMBER: 60/275,578
/ PRIOR FILING DATE: 2001-03-13
/ NUMBER OF SEQ ID NOS: 370
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 74
/ LENGTH: 4675
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-093-463-74
```

Query Match 25.5%; Score 10506; DB 15; Length 4675;

Best Local Similarity 56.1%; Pred. No. 0;

Matches 2252; Conservative 45; Mismatches 129; Indels 1586; Gaps 35;

```
Qy 2595 MPLYNDSFHEISHKGRHTLVLSIQORADAGIVRASSLKVSTSALEVRVPVFLKALD 2654
Db 1 MPLYNDSFHEISHKGRHTLVLSIQORADAGIVRASSLKVSTSALEVRVPVFLKALD 60

Qy 2655 DLSAERTGLALQCEVSDPEAHVVRKDGVLGSDKYDFLHTAGTRGLVVDVSPEDAG 2714
Db 61 DLSAERTGLALQCEVSDPEAHVVRKDGVLGSDKYDFLHTAGTRGLVVDVSPEDAG 120

Qy 2715 LYTHVGSSETRARVRVHDLHGVIKTKLTKMEVLEGESCFECVLSHESADPAMWTVGG 2774
Db 121 LYTHVGSSETRARVRVHDLHGVIKTKLTKMEVLEGESCFECVLSHESADPAMWTVGG 180

Qy 2775 KTVGSSSFQATROGRKYLIVVREAPSDAGEVFSVRGLTSKASLIIVRERPAIIKPLE 2834
Db 181 KTVGSSSFQATROGRKYLIVVREAPSDAGEVFSVRGLTSKASLIIVRERPAIIKPLE 240

Qy 2835 DQWAPGSDVELRCELSRAGTPVHVKDKAIRKSQKDYDVVCEGTMAMLVIRGASLKDAG 2894
Db 241 DQWAPGSDVELRCELSRAGTPVHVKDKAIRKSQKDYDVVCEGTMAMLVIRGASLKDAG 300

Qy 2895 EYTCEVEASKTASLHVKEKANCFTETNLQVEEKGFVFTCKTEHPAAVTVWRKGLLE 2954
Db 301 EYTCEVEASKTASLHVKEKANCFTETNLQVEEKGFVFTCKTEHPAAVTVWRKGLLE 360

Qy 2955 LRASGKHQPSQEGTLRLTITISALEKADSDTYTCDIQGASRAQLLVQGRRVHIIEDLEV 3014
Db 1 LRASGKHQPSQEGTLRLTITISALEKADSDTYTCDIQGASRAQLLVQGRRVHIIEDLEV 420
```

```
Db 361 LRASGKHQPSQEGTLRLTITISALEKADSDTYTCDIQGASRAQLLVQGRRVHIIEDLEV 420
Qy 3015 DVQEGSSATFRCRISpanyEPVHWFLDKTPUHANELNEIDAOPGQYHVLTLRQALAKDSG 3074
Db 421 DVQEGSSATFRCRISpanyEPVHWFLDKTPUHANELNEIDAOPGQYHVLTLRQALAKDSG 480
Qy 3075 TIYFEAGDQORASAAALRVTEKPSVSRELTDATITTEGEDTLIVCETSTCDIPMCWTGDKT 3134
Db 481 TIYFEAGDQORASAAALRVTEKPSVSRELTDATITTEGEDTLIVCETSTCDIPMCWTGDKT 540
Qy 3135 LRGSARCOLSHEGHRAQLLITGATLQDSGRYKCBAGGACSSSIVRVHARPVRFOEALKDL 3194
Db 541 LRGSARCOLSHEGHRAQLLITGATLQDSGRYKCBAGGACSSSIVRVHARPVRFOEALKDL 600
Qy 3195 EVLEGGAAATLRCVLSVVAAPVKWCYGNVLRPGDKYSLRQSGAMLELVVRNLRPODSGRY 3254
Db 601 EVLEGGAAATLRCVLSVVAAPVKWCYGNVLRPGDKYSLRQSGAMLELVVRNLRPODSGRY 660
Qy 3255 SCSFGDQTTSATLTVTALPAQFIGKLRNKEATEGATATLRCELSKTAAPVWRKGSSETLRD 3314
Db 661 SCSFGDQTTSATLTVTALPAQFIGKLRNKEATEGATATLRCELSKTAAPVWRKGSSETLRD 720
Qy 3315 GDRYCLRQDQAMCELOIRGLAMWDAEYSCVCGEERTSASLTIRPMPAHFIRLHRQBSI 3374
Db 721 GDRYCLRQDQAMCELOIRGLAMWDAEYSCVCGEERTSASLTIRPMPAHFIRLHRQBSI 780
Qy 3375 EGATATLRCELSKAAAPVWRKGRSLRDGDRHSRLRQDQAVCELOICGLAVADAGEYSCVC 3434
Db 781 EGATATLRCELSKAAAPVWRKGRSLRDGDRHSRLRQDQAVCELOICGLAVADAGEYSCVC 840
Qy 3435 GEERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLWCELSKVAPVWRKGPENLRDGDY 3494
Db 841 GEERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLWCELSKVAPVWRKGPENLRDGDY 900
Qy 3495 ILROEGTRCELOICGLAMADAGEYLCVCGEERTSATLTIRALPARFIEDVKNQAREGAT 3554
Db 901 ILROEGTRCELOICGLAMADAGEYLCVCGEERTSATLTIRALPARFIEDVKNQAREGAT 960
Qy 3555 AVLQCELSAAPVWRKGSSETLRDGDYSLRQDQTKCELOIRGLAMADTGYSCVCGOER 3614
Db 961 AVLQCELSAAPVWRKGSSETLRDGDYSLRQDQTKCELOIRGLAMADTGYSCVCGOER 1020
Qy 3615 TSAMLTVRALPIKFTTEGLRNEEAEGATAVLRCELSKVAPVWRKGMKHETLRDGRHSLRQ 3674
Db 1021 TSAMLTVRALPIKFTTEGLRNEEAEGATAVLRCELSKVAPVWRKGMKHETLRDGRHSLRQ 1080
Qy 3675 DGARCELOIRGLVAEDAGEYLCMCKERTSAMLTVRAMPKFTIEGLRNEEAEGDTATLW 3734
Db 1081 DGARCELOIRGLVAEDAGEYLCMCKERTSAMLTVRAMPKFTIEGLRNEEAEGDTATLW 1140
Qy 3735 CELSKAAPVWRKGMKHETLRDGRHSLRQDGRSCELOIRGLAVVDVADAGEYSCVCGOERTSAT 3794
Db 1141 CELSKAAPVWRKGMKHETLRDGRHSLRQDGRSCELOIRGLAVVDVADAGEYSCVCGOERTSAT 1200
Qy 3795 LTVRALPARFIEDVKNQAREGATAVLQCELSKAAAPVWRKGSSETLRDGRYSRLRQDGR 3854
Db 1201 LTVRALPARFIEDVKNQAREGATAVLQCELSKAAAPVWRKGSSETLRDGRYSRLRQDGR 1260
Qy 3855 CELQIHGLSVADTGEYSCVCGOERTSATLTVR----- 3886
Db 1261 CELQIHGLSVADTGEYSCVCGOERTSATLTVRALPARFTQDLTKKEASEGATATLQCELS 1320
Qy 3887 ----- 3886
Db 1321 KVAPVWRKGMKHETLRDGRYSRLRQDGRTRCELOIHDLVADAGEYSCVCGOERTSATLTVR 1380
Qy 3887 ----- 3886
Db 1381 ALPARFTEGLRNEEAEGATATLQCELSKAAAPVWRKGMKHETLRDGRYSRLRQDGR 1440
Qy 3887 ----- 3886
Db 1441 IHGLAMADNGVYSCVCGOERTSATLTVRALPARFIEDMRNOKATEGATVTLQCKLRKAAP 1500
```


QY 3887 ----- 3886
Db 1501 VEWKGPNTLKDGDRYSLKQDGTSCELQIRGLVIADAGEYSCICEQRTSATLTVRALPA 1560
QY 3887 ----- 3886
Db 1561 RFIEDVRNHEATEGATAVLOCELSKAAAPVWWRKGSSETLRDGRYSLRQDGTCELOIRGL 1620
QY 3887 ----- 3886
Db 1621 AVEDTGEYLCVCGQERTSATLTVRALPARFIDNMTNQAREGATATLHCELSKAPVWVR 1680
QY 3887 ----- 3886
Db 1681 KGPETLRDGRHSRLRQDGRCELOIRGLAVDAGEYSCVCGQERTSATLTVRALPARFIE 1740
QY 3887 ----- 3886
Db 1741 DVKNQAREGATAVLOCELSKAAAPVWWRKGSSETLRDGRYSLRQDGTCELOIRGLSVAD 1800
QY 3887 ----- 3886
Db 1801 TGEYSCVCGQERTSATLTVRALPARFTQDLTKASEGATATLQCELSKAPVWWRKGSPE 1860
QY 3887 ----- 3886
Db 1861 TLRDGRYSRLKQDGTCELOIRGLHLSVADAGEYSCVCGQERTSATLTVRDCHTLHVMPHY 1920
QY 3887 ----- 3886
Db 1921 POLPGLKEPSETLIYIQIPSPVILFTEGLRNEAMEGATATLQCELSKAAAPVWWRKGLE 1980
QY 3887 ----- 3886
Db 1981 ALRDGDKYSLRDQGAVALCELOIRGLAMADNGVYSSLPARFIEDMRNQKATEGATVTLQCKL 2040
QY 3887 ----- 3886
Db 2041 RKAAPVWWRKGPNTLKDGDRYSRLKQDGTSCELQIRGLVIADAGEYSCICEQRTSATLT 2100
QY 3887 ----- 3886
Db 2101 RALPARFIEDVRNHEATEGATAVLOCELSKAAAPVWWRKGSSETLRDGRYSRLRQDGTCE 2160
QY 3887 ----- 3886
Db 2161 QIRGLAVEDTGEYLCVCGQERTSATLTVRALPARFIDNMTNQAREGATATLHCELSKVA 2220
QY 3887 ----- 3886
Db 2221 PVWWRKGPETLRDGRHSRLRQDGTCELOIRGLSVADAGEYSCVCGQERTSATLTIRALP 2280
QY 3887 ----- 3886
Db 2281 AKFTKGLRNEATEGATAMLOCELSKAPVWWRKGPETLRDGRYLRQDGTCELOIRGL 2340
QY 3887 ----- 3886
Db 2341 LSVADTGEYSCVCGQERTSATLTVRALPARFIDNMTNQAREGATATLHCELSKVA 2400
QY 3922 WSKGLOQLQANGRRPRLOGCTAEVLVQDLQREDTGEYTCGQRTSATLTVAAPVRF 3981
Db 2401 WSKGLOQLQANGRRPRLOGCTAEVLVQDLQREDTGEYTCGQRTSATLTVAAPVRF 2460
QY 3982 LRELQHEVDEGGTAHLCELSRAGASVWWRKGSLOLPFCAPVWWRKGSLOLPFCAPVWWRKGS 4041
Db 2461 LRELQHEVDEGGTAHLCELSRAGASVWWRKGSLOLPFCAPVWWRKGSLOLPFCAPVWWRKGS 2520
QY 4042 QEDAGDYTCDTGHTQSMASLSVRPRPKFTRQLSLEQETGDIARLCCQLSDAESGAVVQ 4101
Db 2521 QEDAGDYTCDTGHTQSMASLSVRGR -----GAA-- 2549

QY 4102 WLKEGVELHAGPKYEMRSQCATRELLIHOLEAKDTGEYACVTGGQKTAASLVRTEPEVTI 4161
Db 2550 -----CGPQVRDAAQAGATRELLIHOLEAKDTGEYACVTGGQKTAASLVRTEPEVTI 2600
QY 4162 VRGLVDAEVTADVEDFEFSCEVSBRAGATGVQWCLQPLQOSNEVTEVAVRDGRHTLRLKG 4221
Db 2601 VRGLVDAEVTADVEDFEFSCEVSBRAGATGVQWCLQPLQOSNEVTEVAVRDGRHTLRLKG 2660
QY 4222 VTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPIQDVOLSEGQDASFQCLRSRAGQEA 4281
Db 2661 VTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPIQDVOLR ----- 2701
QY 4282 RHALCGVPLQANEMNDITVEQGTLLHLLTKHTLEDAGTVSFHVGTCSSSEALQKVT- 4337
Db 2702 -----GVPLQANEMNDITVEQGTLLHLLTKHTLEDAGTVSFHVGTCSSSEALQKVTAVP 2756
QY 4338 ----- 4337
Db 2757 CLVRGLQNVDPVAGEVATFSCEDGPQSAIVRVDGIFHSLMLSLGLGVADSGTVI PRAGPLV 2816
QY 4338 ----- 4337
Db 2817 STAKLLIKDPVVEVVSAMQDLAVEGGSALLCOYSRPVQATWKMDEREVHTDGRHVI 2876
QY 4338 -----AKNTVVVRGLNVLEALGEGALFE 4360
Db 2877 QDMNVARLTFRPALPCDSGIYSCAAAGTRVVALQVQAKNTVVVRGLNVLEALGEGALFE 2936
QY 4361 COLSOPVAAHTWLLDDPVRTSENAEVVFFPENGRLHLLLLKNLRPODSCRVTFLAGDMV 4420
Db 2937 COLSOPVAAHTWLLDDPVRTSENAEVVFFPENGRLHLLLLKNLRPODSCRVTFLAGDMV 2996
QY 4421 TSAFILTVR-----GWRLEILEPLKNAAVRAGAQAARPTCTLSBAVPVGEASWYINGAA 4472
Db 2997 TSAFILTVRGDCAVLVQGWRLLEILEPLKNAAVRAGAQAARPTCTLSBAVPVGEASWYINGAA 3056
QY 4473 VOPDDSDMTVADGSHOALLLSAQPHHAGEVTFACRDVAASARLTVLGLPPEDEAVV 4532
Db 3057 VOPDDSDMTVADGSHHALLLSAQPHHAGEVTFACRDVAASARLTVLGLPPEDEAVV 3116
QY 4533 AHSHTVTLSSWAPMSDGGGLCGYRVEKCATQWRLCHELVPGEPCVVDGLAPGETY 4592
Db 3117 AHSHTVTLSSWAPMSDGGGLCGYRVEKCATQWRLCHELVPGEPCVVDGLAPGETY 3176
QY 4593 RFRVAAGVGVGAGEPVHLPTVRLAEPKPPVPPQPSAPESRQVAAGEDVSLEVVVAEAG 4652
Db 3177 RFRVAAGVGVGAGEPVHLPTVRL-EPKPPVPPQPSAPESRQVAAGEDVCLSEVVVAEAG 3235
QY 4653 EVIHHKGMERTQPGGRFEVVSQGRQOOLVIKGFTHAEDQGEYHCGLAQSGICPAAATFQVA 4712
Db 3236 EVIHHKGMERTQPGGRFEVVSQGRQOOLVIKGFTHAEDQGEYHCGLAQSGICPAAATFQVA 3295
QY 4713 LSPASVDEAPOPSLPPEAAQEGDLHLLWEALARKRMSREPTLDSISELPBEDGRSQRUP 4772
Db 3296 LSPASVDEAPOPSLPPEAAQEGDLHLLWEALARKRMSREPTLDSISELPBEDGRSQRUP 3355
QY 4773 QEAEEVAPDLSEGYSTADELARTGDADLSHTSSDDDESBRAGTSPSLVTLKAGRPGTSPUA 4832
Db 3356 QEAEEVAPDLSEGYSTADELARTGDADLSHTSSDDDESBRAGTSPSLVTLKAGRPGTSPUA 3415
QY 4833 SKVGAP-----AAPSVPKQOOQEP 4852
Db 3416 SKVSPNLAJCKERFPTPRAGRSLGLFVGADPAFPGSGERSARCTTRCAAPPRESLKRPE- 3474
QY 4853 AAVRPLGLDLSKDLG-----DPSMD-----KAAVKIQ-----AAPK 4884
Db 3475 ASCLP--GAMEAVELARKLQBEATCSICLDFTDPMVTTCGHNFRCACIQLSWEKARGKK 3532
QY 4885 GYKVRK-----EMKQOEGPMFSHTFGDTEAQVGDALRLCEVAVSKADVRARWLKDG 4935
Db 3533 GRRKEKSGPCPECEMSPQNLNPN-----RLTKVAEMA----- 3568
QY 4936 VELTDGRHHHIDQLGDGTCSLLIAGLRADAGCYTCQVSNK-----FGQVTHSACVVVSGS 4991


```
QY 2942 PAATVTRKGLLELRASGHOPSOBGLTLRLTISALEKADSTYTCIDIGQASRAQLLVQ 3001
Db 901 PAATVTRKGLLELRASGHOPSOBGLTLRLTISALEKADSTYTCIDIGQASRAQLLVQ 960
QY 3002 GRRVHIIEDELDVDVQEGSSATFCRIISPANYPEVHWFELDKTPLHANELNEIDAOPGGVH 3061
Db 961 GRRVHIIEDELDVDVQEGSSATFCRIISPANYPEVHWFELDKTPLHANELNEIDAOPGGVH 1020
QY 3062 VLTLRQALKDSTGIYFAGDQASAAALRVTEKPSVFSRELTDAITTEGEDTLVCETST 3121
Db 1021 VLTLRQALKDSTGIYFAGDQASAAALRVTEKPSVFSRELTDAITTEGEDTLVCETST 1080
QY 3122 CDIPWCWTQKGTILGARSARCQLSHEGHRQAQLLITGATLQDSGRYKCEAGGACSSSIVRVH 3181
Db 1081 CDIPWCWTQKGTILGARSARCQLSHEGHRQAQLLITGATLQDSGRYKCEAGGACSSSIVRVH 1140
QY 3182 ARPVRFOBALKDLVLEGGATLRCVLSVAAPVKWCYGNVLRPGDKYSLRQEGAMLEL 3241
Db 1141 ARPVRFOBALKDLVLEGGATLRCVLSVAAPVKWCYGNVLRPGDKYSLRQEGAMLEL 1200
QY 3242 VVRNLRPDQSGRYSCSFGDQTTSATLTVTALPAQFIKLRNKEATEGATATLRCELSKTA 3301
Db 1201 VVRNLRPDQSGRYSCSFGDQTTSATLTVTALPAQFIKLRNKEATEGATATLRCELSKAA 1260
QY 3302 PVEWRKGSSETLRDGRYCLRDGAMCELOIRGLAMVDAABEYSCVCGEERTSASLTIRPMP 3361
Db 1261 PVEWRKGSSETLRDGRYCLRDGAMCELOIRGLAMVDAABEYSCVCGEERTSASLTIRPMP 1320
QY 3362 AHFTGRLRHQSIIEGATATLRCELSKAAPVWRKGRSLRDGDRHSLRDGAVCELOICG 3421
Db 1321 AHFTGRLRHQSIIEGATATLRCELSKAAPVWRKGRSLRDGDRHSLRDGAVCELOICG 1380
QY 3422 LAVADAGEYSCVCGEERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLCELSKVAPVEW 3481
Db 1381 LAVADAGEYSCVCGEERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLCELSKVAPVEW 1440
QY 3482 RKGPENLRDGRYILRQSGTRCELOICGLAMADAGEYLCVCGQERTSATLTVRALPARFI 3541
Db 1441 RKGPENLRDGRYILRQSGTRCELOICGLAMADAGEYLCVCGQERTSATLTVRALPARFI 1500
QY 3542 EDVKNQAREGATAVLQCELSAAAPVWRKGSSETLRDGRYSLRDGTRKCELOIRGLAWA 3601
Db 1501 EDVKNQAREGATAVLQCELSAAAPVWRKGSSETLRDGRYSLRDGTRKCELOIRGLAWA 1560
QY 3602 DTGEYSCVCGQERTSAMLTVRALPIKFTTEGLRNEEATEGATAVLRCELSKAPVWVKGH 3661
Db 1561 DTGEYSCVCGQERTSAMLTVRALPIKFTTEGLRNEEATEGATAVLRCELSKAPVWVKGH 1620
QY 3662 ETLRDGRHSLRDGARGCELOIRGLVAEDDAGEYLCMCKERTSAMLTVRAMPSKPIEGLR 3721
Db 1621 ETLRDGRHSLRDGARGCELOIRGLVAEDDAGEYLCMCKERTSAMLTVRAMPSKPIEGLR 1680
QY 3722 NEEATEGDTATLWCELSKAAAPVWRKGHETLRDGRHSLRDGSRCELOIRGLAVDAGE 3781
Db 1681 NEEATEGDTATLWCELSKAAAPVWRKGHETLRDGRHSLRDGSRCELOIRGLAVDAGE 1740
QY 3782 YSCVCGQERTSATLTVRALPARFIEDVKNQAREGATAVLQCELSKAPVWRKGSSETLR 3841
Db 1741 YSCVCGQERTSATLTVRALPARFIEDVKNQAREGATAVLQCELSKAPVWRKGSSETLR 1800
QY 3842 CGDRYSLRDGTRCELOIRGLSVADTGYSCVCGQERTSATLTVRAPOPVFREPLQSLQA 3901
Db 1801 CGDRYSLRDGTRCELOIRGLSVADTGYSCVCGQERTSATLTVRALPARFTQDLKTEA 1860
QY 3902 EEGSTATLQCELSSEPTATVWVSKGLQIQANGREPRLOCGTAEVLVLODLOREDTGYTC 3961
Db 1861 SEGATATLQCELSK--VAPVWRKGPETLRDGRYSLKQDGTGRCELOIRGLSVADAGEYSC 1919
QY 3962 TCGSOATSATLTVAAPVRLRELQHQVDEGTAHLCCELSRAGASVWRKGSLOLFFC 4021
Db 1920 MCGQERTSAMLTVRALPARFTTEGLRNEEATEGATATLQCELSKA--APVWRKGLEALRDG 1978
```

```
QY 4022 AKYQWQDGAABELLVRGVEQEDAGDYTCDTGHTQSMASLSVRVPRPKFKTRLQSLBOET 4081
Db 1979 DKYSLRQDQAVCELOIHLGAMADNGVSCVCGQERTSATLTVRALPARFIEDMRNOKATE 2038
QY 4082 GDIALRCOLSDAESGAVVQMLKEGVELHAGPKYEMRSOGATRELLIHLQLEAKDGTGEVAC 4141
Db 2039 GATVTLQCKLRKA--APVWRKGPNTLKGDRYSLKQDGTSCELQIRGLVITADAGEVSC 2095
QY 4142 VTGQKTAASLRVTEPEVTIVRGLVDAEVTADEDEVEFSCEVSRAGATGVQMCLOGLPIQS 4201
Db 2096 ICEQERTSATLTVRALPARFIEDVRNHEATEGATAVLQCELSKAAP--VWRKGSSETLRD 2153
QY 4202 NEVTEVAVRDRIRHTLRLKGVTPEDAGTVSHLGNHASSAOLTVRAPEVTITILEPQDVQL 4261
Db 2154 GDYRSLR--QDGTRECELOIRGLAVEDEYLCVCGQERTSATLTVRALPARFIDNMNQEBA 2212
QY 4262 SEGQDASQCRLSRASQOEAWALGGVFLQANENNDITVEQGTILHLLTLHKVTLLEDAGTV 4321
Db 2213 REGATATLHCELSKAAAPVE--WRKGRSLRDGRHSLR--QDGAUCELOICGLAVADAGEY 2269
QY 4322 SFHVGTCSSEAOALKV 4336
Db 2270 SCVCGEERTSATLTV 2284

RESULT 9
US-09-858-664A-2
; Sequence 2, Application US/09858664A
; Patent No. US20020072491A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL000927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-664A-2

Query Match 20.4%; Score 8423; DB 9; Length 1665;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1612; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6356 PPSQVITIEDVQAOTGGTAQFAEIIIEGDPQPSVTWYKDSVOLVDSSTRLSQQOEGTYSILV 6415
Db 53 PPSQVITIEDVQAOTGGTAQFAEIIIEGDPQPSVTWYKDSVOLVDSSTRLSQQOEGTYSILV 112
QY 6416 LRHVASKDAGVYTCLAQNTGGQVLCKAELLVLGGNEDPDSKQSHRRKHLHSPFYEVKEBIG 6475
Db 113 LRHVASKDAGVYTCLAQNTGGQVLCKAELLVLGGNEDPDSKQSHRRKHLHSPFYEVKEBIG 172
QY 6476 RGVEGFVKRVQHKGNKILCAAKFIPLASRTAQAYREDDILAAALSHPLVTLGLDQFETRK 6535
Db 173 RGVEGFVKRVQHKGNKILCAAKFIPLASRTAQAYREDDILAAALSHPLVTLGLDQFETRK 232
QY 6536 TLILILECSSEELDLRLYRKGVVTEAEVKYVIOQLVEGLHYLHSHGVHLHDIKPSNITLM 6595
Db 233 TLILILECSSEELDLRLYRKGVVTEAEVKYVIOQLVEGLHYLHSHGVHLHDIKPSNITLM 292
QY 6596 VHPAREDIKICDFGAQNTIPAELOFQSGYSPFVSPEIIOQNVPVSEASDIWAMGVISYL 6655
Db 293 VHPAREDIKICDFGAQNTIPAELOFQSGYSPFVSPEIIOQNVPVSEASDIWAMGVISYL 352
QY 6656 SLTCSPPAGESDRTALLNVLEGRVSWSSPMAHLSEDAKDFIKATLQARQARSAAC 6715
```


QY 6836 AEASERSTEAPAPPASPEGAGPPAAQGCVPRHVSIRSLFYHQAGSPSEHGALAPGSRHRP 6895
DB 533 AEASERSTEAPAPPASPEGAGPPAAQGCVPRHVSIRSLFYHQAGSPSEHGALAPGSRHRP 592
QY 6896 ARRRHLLKGGYTAGALPGLRPLMBHRVLEBEAAREEQATLLAKAPSPETALRLPASGTH 6955
DB 533 ARRRHLLKGGYTAGALPGLRPLMBHRVLEBEAAREEQATLLAKAPSPETALRLPASGTH 652
QY 6956 LAPGSHSLHSDSPSTPRSSSEACGEAQLPSAPSGGAPIRDMGHPQGSQKLPSTGGHPG 7015
DB 653 LAPGSHSLHSDSPSTPRSSSEACGEAQLPSAPSGGAPIRDMGHPQGSQKLPSTGGHPG 712
QY 7016 TAQPERPSPDSWPGOPAPFCHPKQSGAQOEGSCSPHAPVAPCPGSPFPGSCKEAPLVPS 7075
DB 713 TAQPERPSPDSWPGOPAPFCHPKQSGAQOEGSCSPHAPVAPCPGSPFPGSCKEAPLVPS 772
QY 7076 PFLGQOPAPPAPAKASPLDLSKMGPGDISLQGRPKPGPCSSPGSASQASSQVSSLRVGS 7135
DB 773 PFLGQOPAPPAPAKASPLDLSKMGPGDISLQGRPKPGPCSSPGSASQASSQVSSLRVGS 832
QY 7136 SOVGTEPGSLDAEGWTQAEADLSBSTPTLQRPQOATMRKPSLGGRGYAGVAGTGA 7195
DB 833 SOVGTEPGSLDAEGWTQAEADLSBSTPTLQRPQOATMRKPSLGGRGYAGVAGTGA 892
QY 7196 FCGDAGGMLGQGPMMARIATAWVQSSEBEQBARAESQSEBEQBARAESPLQVSGARVP 7255
DB 893 FCGDAGGMLGQGPMMARIATAWVQSSEBEQBARAESQSEBEQBARAESPLQVSGARVP 952
QY 7256 EVGRAPTSSPEPTPWEDIGVSLQIRDLSDGAADTISLDSIEVDPPAYLNLSDLYDI 7315
DB 953 EVGRAPTSSPEPTPWEDIGVSLQIRDLSDGAADTISLDSIEVDPPAYLNLSDLYDI 1012
QY 7316 KYLPPEFMIFRKVPKSAQPEPPSPMAEELAEFPPTWPGELGPHAGLEITESESDVD 7375
DB 1013 KYLPPEFMIFRKVPKSAQPEPPSPMAEELAEFPPTWPGELGPHAGLEITESESDVD 1072
QY 7376 ALLAEAAVGRKKNSSPSRSLFHFGRHLPLDEPAELGLRERVKASVEHISRLKGRPEG 7435
DB 1073 ALLAEAAVGRKKNSSPSRSLFHFGRHLPLDEPAELGLRERVKASVEHISRLKGRPEG 1132
QY 7436 LEKEGPPRRKPGCLASFRSLGSLKSWDRAPTFLRELSDETIVLGQSVTLACQVSAQAQAAT 7495
DB 1133 LEKEGPPRRKPGCLASFRSLGSLKSWDRAPTFLRELSDETIVLGQSVTLACQVSAQAQAAT 1192
QY 7496 WSKOGAPLESSSRVLISATLKNFQLLTILVVAEDLGVTCSVSNALGVTITTVGLRKA 7555
DB 1193 WSKOGAPLESSSRVLISATLKNFQLLTILVVAEDLGVTCSVSNALGVTITTVGLRKA 1252
QY 7556 RPSSSPCPDIGRVDGVLVWKPVESGPVTYIVQCSLEGGSWTTLASDIPDCCYLTSK 7615
DB 1253 RPSSSPCPDIGRVDGVLVWKPVESGPVTYIVQCSLEGGSWTTLASDIPDCCYLTSK 1312
QY 7616 LSRGTYTTRTACVSKAGMGVYSSPSQVLLGGPSHLASEEBSQGRSAQPLSTKTFAQ 7675
DB 1313 LSRGTYTTRTACVSKAGMGVYSSPSQVLLGGPSHLASEEBSQGRSAQPLSTKTFAQ 1372
QY 7676 TQIQGRFSVVRQCKEASGRALAAKIIPIYHPKDKTAVLREYALKGLRHPLHAQ 7735
DB 1373 TQIQGRFSVVRQCKEASGRALAAKIIPIYHPKDKTAVLREYALKGLRHPLHAQ 1432
QY 7736 LSPRLHLVILELCSGPELLPCLAEARASVSESVKQYLQWLSATQYLNQHTLHLDLSE 7795
DB 1433 LSPRLHLVILELCSGPELLPCLAEARASVSESVKQYLQWLSATQYLNQHTLHLDLSE 1492
QY 7796 NMIITEYNLLKVVDLGNQSLQSEKVLPSDKFQVLETPMPELLSGQAVPOTDWAIGV 7855
DB 1493 NMIITEYNLLKVVDLGNQSLQSEKVLPSDKFQVLETPMPELLSGQAVPOTDWAIGV 1552
QY 7856 TAFIMLSAEYPVSSSEGDQRLQRLKGLVRLSRCYAGLSGGGAVAFRLSTLCAQWGRPCA 7915
DB 1553 TAFIMLSAEYPVSSSEGDQRLQRLKGLVRLSRCYAGLSGGGAVAFRLSTLCAQWGRPCA 1612

QY 7916 SSCLOCPWLTREGPACSRPAPVTFTPTARLRFVVRNREKRALLYKRHNLAQVR 7968
DB 1613 SSCLOCPWLTREGPACSRPAPVTFTPTARLRFVVRNREKRALLYKRHNLAQVR 1665
RESULT 11
US-10-697-263-2
; Sequence 2, Application US/10697263
; Publication No. US20040063142A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP-DIV2
; CURRENT APPLICATION NUMBER: US/10/697,263
; CURRENT FILING DATE: 2003-10-31
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-697-263-2
Query Match 20.4%; Score 8423; DB 15; Length 1665;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1612; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 6356 PPSMOVTIEDVQAOQTGGTAQFAEIIIEGDPQPSVTWKDSVOLVDSTRLSQQOEGTTSYL 6415
DB 53 PPSMOVTIEDVQAOQTGGTAQFAEIIIEGDPQPSVTWKDSVOLVDSTRLSQQOEGTTSYL 112
QY 6416 LHRVASKDAGVYTCLAQNTGGQVLCABELLVLGGDNEPDESKQSHRRKLSHSEYKEETG 6475
DB 113 LHRVASKDAGVYTCLAQNTGGQVLCABELLVLGGDNEPDESKQSHRRKLSHSEYKEETG 172
QY 6476 RGVFGFVRVQHKGNKILCAAKFTPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETRK 6535
DB 173 RGVFGFVRVQHKGNKILCAAKFTPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETRK 232
QY 6536 TLILILELCSSEELLDRLYRKGWVTEAEVKYIQQLVEGLHYLHSHGVHLHDIKPSNIUM 6595
DB 233 TLILILELCSSEELLDRLYRKGWVTEAEVKYIQQLVEGLHYLHSHGVHLHDIKPSNIUM 292
QY 6596 VHPAREDIKICDFGPAQNITPAELQFSQYSGSPFVSPBIIQQNPVSEASDIWAMGVISYL 6655
DB 293 VHPAREDIKICDFGPAQNITPAELQFSQYSGSPFVSPBIIQQNPVSEASDIWAMGVISYL 352
QY 6656 SLTCSPPAGSDRATLNLVLEGRVSWSSPMAAHLSEDAKOFIKATLQAPQAPRPSAAQC 6715
DB 353 SLTCSPPAGSDRATLNLVLEGRVSWSSPMAAHLSEDAKOFIKATLQAPQAPRPSAAQC 412
QY 6716 LSHPMFLKSMAPAEAAHPTNTKQLKELLARSWRORSMSYKSLVWRSIPELLRGPDPSPS 6775
DB 413 LSHPMFLKSMAPAEAAHPTNTKQLKELLARSWRORSMSYKSLVWRSIPELLRGPDPSPS 472
QY 6776 LGVARHLCRDTGGSSSSSSSSSDNELAPFARAKSLPPSPVTHSPLLHPRGFLRPSASLPBE 6835
DB 473 LGVARHLCRDTGGSSSSSSSSSDNELAPFARAKSLPPSPVTHSPLLHPRGFLRPSASLPBE 532
QY 6836 AEASERSTEAPAPPASPEGAGPPAAQGCVPRHVSIRSLFYHQAGSPSEHGALAPGSRHRP 6895
DB 533 AEASERSTEAPAPPASPEGAGPPAAQGCVPRHVSIRSLFYHQAGSPSEHGALAPGSRHRP 592
QY 6896 ARRRHLLKGGYTAGALPGLRPLMBHRVLEBEAAREEQATLLAKAPSPETALRLPASGTH 6955


```
Db 593 ARRHLLKGGYTAGALPGLRPLMBHRVLEEAAREEQATLLAKAPSFETALRLPASGTH 652
Qy 6956 LAPGSHSLSHDPSPTPRPSSACGEAORLPAPGGAPIRDMGHPOGSKQLPSTGGHPG 7015
Db 653 LAPGSHSLSHDPSPTPRPSSACGEAORLPAPGGAPIRDMGHPOGSKQLPSTGGHPG 712
Qy 7016 TAQPERPSPDSWPGQAPAPCHPKQSGAPQEGCSPHPAVACPDPGPPGSGCKEAPLPVSS 7075
Db 713 TAQPERPSPDSWPGQAPAPCHPKQSGAPQEGCSPHPAVACPDPGPPGSGCKEAPLPVSS 772
Qy 7076 PFLGOQAPPAPAKASPLDLSKMGCDISLPGRPKPGPCSSPGSASQASSQVSSLRVGS 7135
Db 773 PFLGOQAPPAPAKASPLDLSKMGCDISLPGRPKPGPCSSPGSASQASSQVSSLRVGS 832
Qy 7136 SOVGTEPGPSLDAEGWTQAEADLSSTPTLQRPQOATWKRFSLGGRGYAGVAGTGTA 7195
Db 833 SOVGTEPGPSLDAEGWTQAEADLSSTPTLQRPQOATWKRFSLGGRGYAGVAGTGTA 892
Qy 7196 FCGDAGGMLGQGPMMARIATAVYSQEEBEQEEARAESEEQEEARAESEPLPQVSARPVP 7255
Db 893 FCGDAGGMLGQGPMMARIATAVYSQEEBEQEEARAESEEQEEARAESEPLPQVSARPVP 952
Qy 7256 EVGRAPTRSSPEPTWEDIGQVLSVQIRDLSDAEADTISLDISEVDPAYLNLSDLVDI 7315
Db 953 EVGRAPTRSSPEPTWEDIGQVLSVQIRDLSDAEADTISLDISEVDPAYLNLSDLVDI 1012
Qy 7316 KYLPPEFMIFRKVPKSAOPEPPSMAEBELAFPEPTWPGELGPHAGLETTESEDDVD 7375
Db 1013 KYLPPEFMIFRKVPKSAOPEPPSMAEBELAFPEPTWPGELGPHAGLETTESEDDVD 1072
Qy 7376 ALLABAAVGRKRWKSSPSRSLPHFPGRLPLDEPAELGLRERVKASVEHISRLKGRPEG 7435
Db 1073 ALLABAAVGRKRWKSSPSRSLPHFPGRLPLDEPAELGLRERVKASVEHISRLKGRPEG 1132
Qy 7436 LEKEGPPRKCLASFRLSGLKSWDRAPTFLRELSDETWVLGQSVTLACQVSAQAPAAQT 7495
Db 1133 LEKEGPPRKCLASFRLSGLKSWDRAPTFLRELSDETWVLGQSVTLACQVSAQAPAAQT 1192
Qy 7496 WSKDGAPESSSRVLISATLKNFOLLTILVVAEDLGVTTCVSNALGTVTITGVLRKAE 7555
Db 1193 WSKDGAPESSSRVLISATLKNFOLLTILVVAEDLGVTTCVSNALGTVTITGVLRKAE 1252
Qy 7556 RPSSSPCPDIEGVYADGVLVWKPVEYGPVTYIVQCSLEGSSWTLASDIFDCCYLTSK 7615
Db 1253 RPSSSPCPDIEGVYADGVLVWKPVEYGPVTYIVQCSLEGSSWTLASDIFDCCYLTSK 1312
Qy 7616 LSRGTYTFTACVSKAGWGPVSSPEOVLLGGPSHLASEESQGRSAQPLPSTKTFAQ 7675
Db 1313 LSRGTYTFTACVSKAGWGPVSSPEOVLLGGPSHLASEESQGRSAQPLPSTKTFAQ 1372
Qy 7676 TQIQRGFRSVVRQCWEKASGRALAAKIIPYHPKDKTAVLREYEAALKGLRHPHQAQHAAY 7735
Db 1373 TQIQRGFRSVVRQCWEKASGRALAAKIIPYHPKDKTAVLREYEAALKGLRHPHQAQHAAY 1432
Qy 7736 LSPRHLVILELCSPPELLPCLAEARASYSESEVKDYLWQMLSATQYLHNQHILHLDLRSE 7795
Db 1433 LSPRHLVILELCSPPELLPCLAEARASYSESEVKDYLWQMLSATQYLHNQHILHLDLRSE 1492
Qy 7796 NMIIETEYLLKVVDLIGNAQSLSOEKLPSDKFKDYLETWAPELLEGGQAVPQTDIWAIGV 7855
Db 1493 NMIIETEYLLKVVDLIGNAQSLSOEKLPSDKFKDYLETWAPELLEGGQAVPQTDIWAIGV 1552
Qy 7856 TAFIMLSAEYPVSSSGARDLQRLKGLVRLSRVYAGLSGGAVAFRLSTLCAQWGRPCA 7915
Db 1553 TAFIMLSAEYPVSSSGARDLQRLKGLVRLSRVYAGLSGGAVAFRLSTLCAQWGRPCA 1612
Qy 7916 SSCLOCPWLTEBGPACSRPAPVTPTTARLRFVVRNREKRALLYKRHNLAQVR 7968
Db 1613 SSCLOCPWLTEBGPACSRPAPVTPTTARLRFVVRNREKRALLYKRHNLAQVR 1665
```

RESULT 12

US-10-921-168-2

```
; Sequence 2, Application US/10921168
; Publication No. US20050003446A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: THEREOF
; FILE REFERENCE: CLO00927-CIP-DIV3
; CURRENT APPLICATION NUMBER: US/10/921,168
; CURRENT FILING DATE: 2004-08-19
; PRIOR APPLICATION NUMBER: 10/274,978
; PRIOR FILING DATE: 2002-10-22
; PRIOR APPLICATION NUMBER: 09/858,664
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1665
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-921-168-2
```

```
Query Match 20.4%; Score 8423; DB 16; Length 1665;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1612; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 6356 PPSQVITIEDVOAQTGCTAQFEAIIEGDPQSPVTWYKDSVOLVDSTRLSQQOEGTYSVLV 6415
Db 53 PPSQVITIEDVOAQTGCTAQFEAIIEGDPQSPVTWYKDSVOLVDSTRLSQQOEGTYSVLV 112
Qy 6416 LRHVASKDAGVYTCCLAQNTGGQVLCKAELLVGLGDNEDPDSKQSHRRKLHGFYEVKEIG 6475
Db 113 LRHVASKDAGVYTCCLAQNTGGQVLCKAELLVGLGDNEDPDSKQSHRRKLHGFYEVKEIG 172
Qy 6476 RGVEFGVKRVQHKNGKILCAAKFTPLRSRTAAQYRERDILAAALSHPLVTGLLQGFETR 6535
Db 173 RGVEFGVKRVQHKNGKILCAAKFTPLRSRTAAQYRERDILAAALSHPLVTGLLQGFETR 232
Qy 6536 TLILILELCSSEELLDRLYRKGVVTEAEVKYIIOOLVEGLHYLHSHGVHLHDIKPSNLLM 6595
Db 233 TLILILELCSSEELLDRLYRKGVVTEAEVKYIIOOLVEGLHYLHSHGVHLHDIKPSNLLM 292
Qy 6596 VHPAREDIKICDFGAQNIITPAELQFSQYSPFVSPPEIIQONPVSEASDIWAMGVISYL 6655
Db 293 VHPAREDIKICDFGAQNIITPAELQFSQYSPFVSPPEIIQONPVSEASDIWAMGVISYL 352
Qy 6656 SLTCSPPFAGESDRATLLNVLEGRVSWSSPMAAHLSEDAKDFIKATLQAPQARPSAAQC 6715
Db 353 SLTCSPPFAGESDRATLLNVLEGRVSWSSPMAAHLSEDAKDFIKATLQAPQARPSAAQC 412
Qy 6716 LSHPWFLKSPMAEBAHINTKQLKFLARSRWQSLMSYKSLVMRSITPELLRGPDPSPS 6775
Db 413 LSHPWFLKSPMAEBAHINTKQLKFLARSRWQSLMSYKSLVMRSITPELLRGPDPSPS 472
Qy 6776 LGVARHLCRDTGGSSSSSSSDNDELAPFARAKSLPPSPVTHSPLLHPRGFLRPSASLPEE 6835
Db 473 LGVARHLCRDTGGSSSSSSSDNDELAPFARAKSLPPSPVTHSPLLHPRGFLRPSASLPEE 532
Qy 6836 AEASERSTEAPAPASPEGAGPPAAQOCVPRHVSIRSLFYHQAGESPEHGALAPGSRHRP 6895
Db 533 AEASERSTEAPAPASPEGAGPPAAQOCVPRHVSIRSLFYHQAGESPEHGALAPGSRHRP 592
Qy 6896 ARRHLLKGGYIAGALPGLREPIMEHVLREEAAREEQATLLAKAPSFETALRLPASGTH 6955
Db 593 ARRHLLKGGYIAGALPGLREPIMEHVLREEAAREEQATLLAKAPSFETALRLPASGTH 652
Qy 6956 LAPGSHSLSHDPSPTPRPSSACGEAORLPAPGGAPIRDMGHPOGSKQLPSTGGHPG 7015
Db 653 LAPGSHSLSHDPSPTPRPSSACGEAORLPAPGGAPIRDMGHPOGSKQLPSTGGHPG 712
Qy 7016 TAQPERPSPDSWPGQAPAPCHPKQSGAPQEGCSPHPAVACPDPGPPGSGCKEAPLPVSS 7075
```



```
Db 713 TAQPERSPDSWGGQAPCPCHPKQASAPQGGCSHPAVAPCPGSPFPQGSCKEADLVPS 772
Qy 7076 PFLGQOPAPAPAKASPLDSSXMGPCDLSLPGKPKPGPCSSPGSASQASSSSQVSLRVGS 7135
Db 773 PFLGQOPAPAPAKASPLDSSXMGPCDLSLPGKPKPGPCSSPGSASQASSSSQVSLRVGS 832
Qy 7136 SOVGTEPGSLDAEGTQAEADLSSTPTLPQBPQATNRKESLGGRGYAGVAGYGFPA 7195
Db 833 SOVGTEPGSLDAEGTQAEADLSSTPTLPQBPQATNRKESLGGRGYAGVAGYGFPA 892
Qy 7196 FCGDAGGMLGQPMWARIAWVSQSEEEQSEARAESEEEQSEARAESEPLQVQARVPV 7255
Db 893 FCGDAGGMLGQPMWARIAWVSQSEEEQSEARAESEEEQSEARAESEPLQVQARVPV 952
Qy 7256 EVGRAPTRSSPEPTWEDIGQVSLVQIRDLSDGAADTISLDSIEVDPAAYLNLSLDYDI 7315
Db 953 EVGRAPTRSSPEPTWEDIGQVSLVQIRDLSDGAADTISLDSIEVDPAAYLNLSLDYDI 1012
Qy 7316 KYLPFERMI FRKVPKSAQPEPPMAEELAEFPPTWPGELGPHAGLEITSESDVD 7375
Db 1013 KYLPFERMI FRKVPKSAQPEPPMAEELAEFPPTWPGELGPHAGLEITSESDVD 1072
Qy 7376 ALLAAAVCRKRWSSPKRSLFHPGRHLPLDEPAELGLRERVKASVEHISRLKGRPEG 7435
Db 1073 ALLAAAVCRKRWSSPKRSLFHPGRHLPLDEPAELGLRERVKASVEHISRLKGRPEG 1132
Qy 7436 LEKEGPRKKGLASFRLSGLKSWDRAPTFLRELSDETIVLQGSVTLACQVSAQPAQAAT 7495
Db 1133 LEKEGPRKKGLASFRLSGLKSWDRAPTFLRELSDETIVLQGSVTLACQVSAQPAQAAT 1192
Qy 7496 WSKGAPLESSSRVLSATLKNFOLLTLVVAEDLGVYTCVSNALGTVTTTGVLRKAE 7555
Db 1193 WSKGAPLESSSRVLSATLKNFOLLTLVVAEDLGVYTCVSNALGTVTTTGVLRKAE 1252
Qy 7556 RPSSPCPDIGVYADGVLLVWKPVSYPVTYIVQCSLEGGSWTTLASDIFDCCVLTSK 7615
Db 1253 RPSSPCPDIGVYADGVLLVWKPVSYPVTYIVQCSLEGGSWTTLASDIFDCCVLTSK 1312
Qy 7616 LSRGTYTFRACVSKAGMGYPSSPSEQVLLGGPSHLSAEESQGRSAQPLPSTKTFAPQ 7675
Db 1313 LSRGTYTFRACVSKAGMGYPSSPSEQVLLGGPSHLSAEESQGRSAQPLPSTKTFAPQ 1372
Qy 7676 TOIQGRSVVROCKEKASGRALAAKIIPYHPKDKTAVLREYEAALKGLRPHLAQHAAY 7735
Db 1373 TOIQGRSVVROCKEKASGRALAAKIIPYHPKDKTAVLREYEAALKGLRPHLAQHAAY 1432
Qy 7736 LSPRHLVILRLCSGPPELLPCLAEARASYSESEVKDYLWQMLSATQYLHNQHLHDLRSE 7795
Db 1433 LSPRHLVILRLCSGPPELLPCLAEARASYSESEVKDYLWQMLSATQYLHNQHLHDLRSE 1492
Qy 7796 NMITEYNLLKYVDLGNQSLSQEKVLPSPDKFYLETWAPELLBGQGAVPQTDIWAIGV 7855
Db 1493 NMITEYNLLKYVDLGNQSLSQEKVLPSPDKFYLETWAPELLBGQGAVPQTDIWAIGV 1552
Qy 7856 TAFIMLSAEYPVSSGARDLQRLKGLVRLSRCVAGLGGGAVAFRLSTLCAQWGRPCA 7915
Db 1553 TAFIMLSAEYPVSSGARDLQRLKGLVRLSRCVAGLGGGAVAFRLSTLCAQWGRPCA 1612
Qy 7916 SSCLOCPMLTEGPACSRPAPVTFPTARLVFVRNREKERRALLYKRNLAQVR 7968
Db 1613 SSCLOCPMLTEGPACSRPAPVTFPTARLVFVRNREKERRALLYKRNLAQVR 1665
```

RESULT 13

US-10-415-011-22

```
; Sequence 22, Application US/10415011
; Publication No. US20040053394A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION
; APPLICANT: GURURAJAN, Rajagopal
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: CHAWLA, Narinder K.
```

```
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: XU, Yuming
; APPLICANT: ARVIZU, Chandra S.
; APPLICANT: YAO, Monique G.
; APPLICANT: RAMKUMAR, Jayalaxmi
; APPLICANT: DING, Li
; APPLICANT: TANG, Y. Tom
; APPLICANT: HAFALIA, April J.A.
; APPLICANT: NGUYEN, Dannie B.
; APPLICANT: GANDHI, Ameena R.
; APPLICANT: LU, Yan
; APPLICANT: YUE, Henry
; APPLICANT: BUREFORD, Neil
; APPLICANT: BANDMAN, Olga
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: LAL, Preeti G.
; APPLICANT: RECIPON, Shirley A.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: BOROWSKY, Mark L.
; APPLICANT: THORNTON, Michael B.
; APPLICANT: SWARNAKER, Anita
; APPLICANT: THANGAVELU, Kavitha
; APPLICANT: KHAN, Farrah A.
; APPLICANT: ISON, Craig H.
; TITLE OF INVENTION: HUMAN KINASES
; FILE REFERENCE: PI-0262 USN
; CURRENT APPLICATION NUMBER: US/10/415,011
; CURRENT FILING DATE: 2003-04-18
; PRIOR APPLICATION NUMBER: PCT/US01/47728
; PRIOR FILING DATE: 2001-10-20
; PRIOR APPLICATION NUMBER: US 60/242,410
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: US 60/244,068
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: US 60/245,708
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/247,672
; PRIOR FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: US 60/249,565
; PRIOR FILING DATE: 2000-11-16
; PRIOR APPLICATION NUMBER: US 60/252,730
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: US 60/250,807
; PRIOR FILING DATE: 2000-12-01
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: PERL Program
; SEQ ID NO 22
; LENGTH: 1665
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20040053394A1 7638121CDI
US-10-415-011-22
```

```
Query Match 20.4%; Score 9410; DB 15; Length 1665;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1610; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 6356 PPSMQVTIEDVQAQTGGTAQFEAIIIEGDPQPSVTHYKDSVOLVDSRLSQOQEGTYSIV 6415
Db 53 PPSMQVTIEDVQAQTGGTAQFEAIIIEGDPQPSVTHYKDSVOLVDSRLSQOQEGTYSIV 112
Qy 6416 LRHVASKDAGVYVTCIAQNTGGQVLCCKAELLVLLGGNEPDSKQSHRRKLHSFYVKEEIG 6475
Db 113 LRHVASKDAGVYVTCIAQNTGGQVLCCKAELLVLLGGNEPDSKQSHRRKLHSFYVKEEIG 172
Qy 6476 RGVFGFVKRVQHKGNKILCAAKFTPLRSRTRAQVREDDILAAALSHPLVTGLLDQFETRK 6535
Db 173 RGVFGFVKRVQHKGNKILCAAKFTPLRSRTRAQVREDDILAAALSHPLVTGLLDQFETRK 232
Qy 6536 TLILILECSSEELLRLRYKRGVVTAEVKYIQQLVESGLHYLHSHGVHLHLDIKPSNIIIM 6595
```


Db 233 TLILILELCSBELDLRLYRKGVVTEAEVKYIQQLVEGLHYLHSHGVHLHLDIKPSNIIM 292
Qy 6596 VHPAREDIKICDFGAQNTTPAELQFSQYSGSEFVSPEIIQQNPVSEASDIWAMGVISYL 6655
Db 293 VHPAREDIKICDFGAQNTTPAELQFSQYSGSEFVSPEIIQQNPVSEASDIWAMGVISYL 352
Qy 6656 SUTCSSPPAGESDRATLLNLVLEGYSWSSPMAAHLSEDAKOFIKATLQAPQARPSAAQC 6715
Db 353 SUTCSSPPAGESDRATLLNLVLEGYSWSSPMAAHLSEDAKOFIKATLQAPQARPSAAQC 412
Qy 6716 LSHPWFLKSMPEABEAFINTKQLKLLARSWQSRMSYKSLVWRSIPELLRGPDSPS 6775
Db 413 LSHPWFLKSMPEABEAFINTKQLKLLARSWQSRMSYKSLVWRSIPELLRGPDSPS 472
Qy 6776 LGVAHRLCRDTCGSSSSSSSSSNEELAPPARAASLPPSPVTHSPLLHPGFLRPSASLPEE 6835
Db 473 LGVAHRLCRDTCGSSSSSSSSSNEELAPPARAASLPPSPVTHSPLLHPGFLRPSASLPEE 532
Qy 6836 AEASERSTEAPAPPASPEGAGPPAAQGCVPRHVSIRSLFYHOAGBSPHEGALAPGSRHRP 6895
Db 533 AEASERSTEAPAPPASPEGAGPPAAQGCVPRHVSIRSLFYHOAGBSPHEGALAPGSRHRP 592
Qy 6896 ARRRHLKGGYTAGALPGLREPLMEHRVLEEAAREEQATLLAKAPSPETALRLPASGTH 6955
Db 593 ARRRHLKGGYTAGALPGLREPLMEHRVLEEAAREEQATLLAKAPSPETALRLPASGTH 652
Qy 6956 LAPGHSHLEHDSPTPRPSSACGEAORLPAPSGGAPIRDMGHPOGSKOLPSTGGHPG 7015
Db 653 LAPGHSHLEHDSPTPRPSSACGEAORLPAPSGGAPIRDMGHPOGSKOLPSTGGHPG 712
Qy 7016 TAQPERPSPDPSWGPAPPCHPKQGSAPQEGCSPPHAPVACPFGSPFPFGSCKEAPLVSS 7075
Db 713 TAQPERPSPDPSWGPAPPCHPKQGSAPQEGCSPPHAPVACPFGSPFPFGSCKEAPLVSS 772
Qy 7076 PFLGQOQAPPAPAKASPLDLSKMGCDISLPGRPXPGCCSPGCSASQASSQVSSLRVS 7135
Db 773 PFLGQOQAPPAPAKASPLDLSKMGCDISLPGRPXPGCCSPGCSASQASSQVSSLRVS 832
Qy 7136 SOVGTPEGPSLDAEGWTQAEADLSDTPTLQRPQEQATMRKFSLGGRGYAGVAGYGTFA 7195
Db 833 SOVGTPEGPSLDAEGWTQAEADLSDTPTLQRPQEQATMRKFSLGGRGYAGVAGYGTFA 892
Qy 7196 FCGDAGGMLGQGPMMARIATAVSOSEBEEQEPARAESQSEEQEAREASPLPQVSARVP 7255
Db 893 FCGDAGGMLGQGPMMARIATAVSOSEBEEQEPARAESQSEEQEAREASPLPQVSARVP 952
Qy 7256 EVGRAPTRSSPEPTPWEDIGQVSLVQIIRDLSDGDAEAADTISLDISEVDPAYNLSDLYDI 7315
Db 953 EVGRAPTRSSPEPTPWEDIGQVSLVQIIRDLSDGDAEAADTISLDISEVDPAYNLSDLYDI 1012
Qy 7316 KYLPPEFMI FRKVPKSAQPEPPSMAEELAEFPPTWPGELGPHAGLETTESEEDVD 7375
Db 1013 KYLPPEFMI FRKVPKSAQPEPPSMAEELAEFPPTWPGELGPHAGLETTESEEDVD 1072
Qy 7376 ALLAAEAAGVRKRKWSPPSRSLFHPFGRLPLDEPAELGLRERVKASVEHISRIILKGRPEG 7435
Db 1073 ALLAAEAAGVRKRKWSPPSRSLFHPFGRLPLDEPAELGLRERVKASVEHISRIILKGRPEG 1132
Qy 7436 LEKEGPPRRKPCGLASFRISGLKSWDRAPTFILRELSDETIVLQSVTLACQVSAQAPAAQ 7495
Db 1133 LEKEGPPRRKPCGLASFRISGLKSWDRAPTFILRELSDETIVLQSVTLACQVSAQAPAAQ 1192
Qy 7496 WSKGAPLESSRVLISATLKNFOLLTILVVAEDLVYTVSVSNALGTVTITGVLRKAE 7555
Db 1193 WSKGAPLESSRVLISATLKNFOLLTILVVAEDLVYTVSVSNALGTVTITGVLRKAE 1252
Qy 7556 RPSSSPCPDIEGVYADGVLLVWKPVSYPVTYIVQCSLEGGSWTLASDIFDCCYLTSK 7615
Db 1253 RPSSSPCPDIEGVYADGVLLVWKPVSYPVTYIVQCSLEGGSWTLASDIFDCCYLTSK 1312
Qy 7616 LSRGCTYTRTACVSKAGMPYSSPEQVTLGGPSHLASEESQGRSAQPLPSTKTFAQ 7675
Db 1313 LSRGCTYTRTACVSKAGMPYSSPEQVTLGGPSHLASEESQGRSAQPLPSTKTFAQ 1372

Qy 7676 TOIQGRFSVVRQCEKASGRALAAKIIPYHPKDKTAVLREYEALKGLRHPHLAOLHAAY 7735
Db 1373 TOIQGRFSVVRQCEKASGRALAAKIIPYHPKDKTAVLREYEALKGLRHPHLAOLHAAY 1432
Qy 7736 LSPRHLVILILELCSGPELLPCLABRASYSSEVKDYLMQMLSATQYLNQHLHLDLRSE 7795
Db 1433 LSPRHLVILILELCSGPELLPCLABRASYSSEVKDYLMQMLSATQYLNQHLHLDLRSE 1492
Qy 7796 NMIIITEYNLLKVDLGNASQSLQSKVLPSPDKFKDYLETMAPELLEGQAGVPTDIWATGV 7855
Db 1493 NMIIITEYNLLKVDLGNASQSLQSKVLPSPDKFKDYLETMAPELLEGQAGVPTDIWATGV 1552
Qy 7856 TAFIMLSAEYPVSSSEGARDLQRLKGLVRLSRVCYAGLSGGAVAFRLSTLCAQPWGRPCA 7915
Db 1553 TAFIMLSAEYPVSSSEGARDLQRLKGLVRLSRVCYAGLSGGAVAFRLSTLCAQPWGRPCA 1612
Qy 7916 SSCLQCPWLTEBGPACSRPAPVPTPTARLRFVVRNREKRALLYKRHNLAQVR 7968
Db 1613 SSCLQCPWLTEBGPACSRPAPVPTPTARLRFVVRNREKRALLYKRHNLAQVR 1665

RESULT 14

US-10-182-243-46
; Sequence 46, Application US/10182243
; Publication No. US20040048310A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY D.
; APPLICANT: WHITE, DAVID
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; APPLICANT: MARTINEZ, RICARDO
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE
; TITLE OF INVENTION: ENZYMS
; FILE REFERENCE: 038602/1366
; CURRENT APPLICATION NUMBER: US/10/182,243
; CURRENT FILING DATE: 2003-07-07
; PRIOR APPLICATION NUMBER: PCT/US01/02337
; PRIOR FILING DATE: 2001-01-25
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 46
; LENGTH: 1618
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-182-243-46

Query Match 20.4%; Score 8407; DB 15; Length 1618;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1612; Conservative 0; Mismatches 0; Indels 6; Gaps 1;

Qy 6357 PSMQVTTIEDVQAQGTGTAQFEAIILEGDPQPSVTHYKDSVOLVDSTRLSQOEGTYSVL 6416
Db 1 PSMQVTTIEDVQAQGTGTAQFEAIILEGDPQPSVTHYKDSVOLVDSTRLSQOEGTYSVL 60
Qy 6417 RHVASKDAGVYTCLAQNTGGQVLCKAEILLVLG-----GDNEPDSEKQSHRKLHSFYE 6470
Db 61 RHVASKDAGVYTCLAQNTGGQVLCKAEILLVLGASHLSIGDNEPDSEKQSHRKLHSFYE 120
Qy 6471 KEEIGRGVFGVKRVQHGKNIKCAAKFIPLRSTRQAQYRERDILAAALSHPLVTGLDQ 6530
Db 121 KEEIGRGVFGVKRVQHGKNIKCAAKFIPLRSTRQAQYRERDILAAALSHPLVTGLDQ 180
Qy 6531 PETRKTLLILELCSBELDLRLYRKGVVTEAEVKYIQQLVEGLHYLHSHGVHLHLDIKP 6590
Db 181 PETRKTLLILELCSBELDLRLYRKGVVTEAEVKYIQQLVEGLHYLHSHGVHLHLDIKP 240
Qy 6591 SNILMVHPAREDIKICDFGAQNTTPAELQFSQYSGSEFVSPEIIQQNPVSEASDIWAMG 6650
Db 241 SNILMVHPAREDIKICDFGAQNTTPAELQFSQYSGSEFVSPEIIQQNPVSEASDIWAMG 300
Qy 6651 VISVLSLTCSSPPAGESDRATLLNLVLEGYSWSSPMAAHLSEDAKOFIKATLQAPQARP 6710

QY	6839	SERSTEAPAPPASPEGAGPPAAQGCVPVRSVIRSLFYHQGESPBHGALAPGSRHRPARR	6898
Db	481	SERSTEAPAPPASPEGAGPPAAQGCVPVRSVIRSLFYHQGESPBHGALAPGSRHRPARR	540
QY	6899	RHLKGGYTAGALPGLREPMHRVLEBEAAREBQATLLAKAPSFETALRLPASGTHLAP	6958
Db	541	RHLKGGYTAGALPGLREPMHRVLEBEAAREBQATLLAKAPSFETALRLPASGTHLAP	600
QY	6959	GHSLSLEHDSPTPRPSSACGEAORLPSAPSGGAPIRDMGHPQSGKOLPSTGGHPGTAQ	7018
Db	601	GHSLSLEHDSPTPRPSSACGEAORLPSAPSGGAPIRDMGHPQSGKOLPSTGGHPGTAQ	660
QY	7019	PERPSPDSWPQPAFCHPQKQSAQEGCSHPVAPCPGSPFPGSCKEAPLVPSSPFL	7078
Db	661	PERPSPDSWPQPAFCHPQKQSAQEGCSHPVAPCPGSPFPGSCKEAPLVPSSPFL	720
QY	7079	GQFQAPPAPAKASPLDLSKMGPGDISLPGRPKPGPCSPGSSAQASSQVSSRLRVGSSQV	7138
Db	721	GQFQAPPAPAKASPLDLSKMGPGDISLPGRPKPGPCSPGSSAQASSQVSSRLRVGSSQV	780
QY	7139	GTEPGPSLDAEGWTQEAEDLSSTPTLQRPQATMRKFSLGGRGYAGVAGYGTFAFGG	7198
Db	781	GTEPGPSLDAEGWTQEAEDLSSTPTLQRPQATMRKFSLGGRGYAGVAGYGTFAFGG	840
QY	7199	DAGGMLGQGPMMARIAMAVSQSEEBEQAERAESEQCEARAESEPLQVSARVPVEVG	7258
Db	841	DAGGMLGQGPMMARIAMAVSQSEEBEQAERAESEQCEARAESEPLQVSARVPVEVG	900
QY	7259	RAPTRSSPEPTWEDIGQVSLVQIRDLSGDAEADTISLDISEVDPAYLNISDLYDIKYL	7318
Db	901	RAPTRSSPEPTWEDIGQVSLVQIRDLSGDAEADTISLDISEVDPAYLNISDLYDIKYL	960
QY	7319	PPEFMI FRKVPKSAQPEPPSMAEBELAEFPPTWPMGELGPHAGLEITBESEDDVALL	7378
Db	961	PPEFMI FRKVPKSAQPEPPSMAEBELAEFPPTWPMGELGPHAGLEITBESEDDVALL	1020
QY	7379	ABAAVGRKRWSSPSRSLFHPFGRHLPLDEPAELGLRERVKASVEHISRILKGRPEGLEK	7438
Db	1021	ABAAVGRKRWSSPSRSLFHPFGRHLPLDEPAELGLRERVKASVEHISRILKGRPEGLEK	1080
QY	7439	EGPRPKKCLASFRISGLKSWDRAPFLREISDETVVLGQSVTLACQVSAQPAQAATWSK	7498
Db	1081	EGPRPKKCLASFRISGLKSWDRAPFLREISDETVVLGQSVTLACQVSAQPAQAATWSK	1140
QY	7499	DCAPLESSRVLISATLKNFOLLITLVVVAEDLVYTCVSNALGTVTTTGVLRKAERPS	7558
Db	1141	DCAPLESSRVLISATLKNFOLLITLVVVAEDLVYTCVSNALGTVTTTGVLRKAERPS	1200
QY	7559	SSPCPDIGEVVADGVLLVWKPVESYGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSR	7618
Db	1201	SSPCPDIGEVVADGVLLVWKPVESYGPVTYIVQCSLEGGSWTTLASDIFDCCYLTSKLSR	1260
QY	7619	GGTYTFRCTACVKAGMGYSFSPSEOVTLGGPSHLASESESQGRSAQPLPSTKTFAFQOI	7678
Db	1261	GGTYTFRCTACVKAGMGYSFSPSEOVTLGGPSHLASESESQGRSAQPLPSTKTFAFQOI	1320
QY	7679	QGRSVVVRQWEKASGRALAAKIIPYHPKOKTAVLREYEAALKGLRHPHLAQLHAAVLSP	7738
Db	1321	QGRSVVVRQWEKASGRALAAKIIPYHPKOKTAVLREYEAALKGLRHPHLAQLHAAVLSP	1380
QY	7739	RHLVLIILELCSGPELLPCLAEASYESSEVKDYLMQMLSATQYLHNQHILHDLRSENMI	7798
Db	1381	RHLVLIILELCSGPELLPCLAEASYESSEVKDYLMQMLSATQYLHNQHILHDLRSENMI	1440
QY	7799	ITEYNLLKVVLDLGNQASLSQEKVLPSPDKPQLETMAPELLEGGAVPQTDIWAIGVTAF	7858
Db	1441	ITEYNLLKVVLDLGNQASLSQEKVLPSPDKPQLETMAPELLEGGAVPQTDIWAIGVTAF	1500
QY	7859	IMLSAEYFVSSSEGARDLQRLRKGIVRLSRVCYAGLSGGAVAFRLRSTLCAQPWGRPCASSC	7918
Db	1501	IMLSAEYFVSSSEGARDLQRLRKGIVRLSRVCYAGLSGGAVAFRLRSTLCAQPWGRPCASSC	1560

QY	7919	LQCPWLTEEGPACSRPAPVTPTARLRFVFNREKRALLYKRHNLAQVR	7968
Db	1561	LQCPWLTEEGPACSRPAPVTPTARLRFVFNREKRALLYKRHNLAQVR	1610

Search completed: March 18, 2005, 18:05:29
Job time : 554 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 17:32:23 ; Search time 126 Seconds
(without alignments)
4720.665 Million cell updates/sec

Title: US-10-077-130-5
Perfect score: 41273
Sequence: 1 MDQPFSGAPFLRTPKAFV.....RREKRALLYRHNLAQVR 7968

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	8423	20.4	1665	4	US-09-858-664A-2
2	8423	20.4	1665	4	US-10-274-978-2
3	8423	20.4	1665	4	US-10-697-263-2
4	1173	2.8	846	4	US-09-858-664A-3
5	1173	2.8	846	4	US-10-274-978-4
6	1173	2.8	846	4	US-10-697-263-4
7	1074	2.6	4391	4	US-10-006-011A-2
8	1025	2.5	1953	4	US-09-917-254-92
9	909.5	2.2	549	4	US-09-858-664A-5
10	909.5	2.2	549	4	US-10-274-978-6
11	909.5	2.2	549	4	US-10-697-263-6
12	892.5	2.2	2860	2	US-08-826-267-2
13	592.5	1.4	414	4	US-09-858-664A-13
14	592.5	1.4	414	4	US-10-274-978-14
15	592.5	1.4	414	4	US-10-697-263-14
16	572.5	1.4	1709	4	US-09-949-016-10503
17	542.5	1.3	615	4	US-09-949-016-11320
18	531	1.3	279	4	US-09-858-664A-4
19	531	1.3	279	4	US-10-274-978-5
20	531	1.3	279	4	US-10-697-263-5
21	521	1.3	298	4	US-09-858-664A-17
22	521	1.3	298	4	US-10-274-978-18
23	521	1.3	298	4	US-10-697-263-18
24	511.5	1.2	508	4	US-09-858-664A-18
25	511.5	1.2	508	4	US-10-274-978-19
26	511.5	1.2	508	4	US-10-697-263-19
27	504	1.2	250	4	US-09-858-664A-6

28	504	1.2	250	4	US-10-274-978-7	Sequence 7, Appl1
29	504	1.2	250	4	US-10-697-263-7	Sequence 7, Appl1
30	497.5	1.2	596	4	US-09-797-039-8	Sequence 8, Appl1
31	495	1.2	969	4	US-09-949-016-8059	Sequence 8059, Ap
32	489	1.2	448	2	US-09-159-385-2	Sequence 2, Appl1
33	489	1.2	448	3	US-09-186-277-2	Sequence 2, Appl1
34	476.5	1.2	11877	3	US-09-105-537-6	Sequence 6, Appl1
35	472.5	1.1	260	2	US-07-857-224B-23	Sequence 23, Appl
36	471	1.1	1431	4	US-09-538-092-1198	Sequence 1198, Ap
37	470	1.1	502	4	US-09-949-016-7837	Sequence 7837, Ap
38	466.5	1.1	454	2	US-09-159-385-1	Sequence 1, Appl1
39	466.5	1.1	454	3	US-09-186-277-1	Sequence 1, Appl1
40	466.5	1.1	454	4	US-09-949-016-6066	Sequence 6066, Ap
41	466	1.1	274	4	US-09-858-664A-14	Sequence 14, Appl
42	466	1.1	274	4	US-10-274-978-15	Sequence 15, Appl
43	466	1.1	274	4	US-10-697-263-15	Sequence 15, Appl
44	465	1.1	358	4	US-09-230-896C-29	Sequence 29, Appl
45	464.5	1.1	1050	3	US-09-428-711A-16	Sequence 16, Appl

ALIGNMENTS

RESULT 1
US-09-858-664A-2
; Sequence 2, Application US/09858664A
; Patent No. 6482624
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; CURRENT FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; PRIOR FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1665
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-664A-2

Query Match	20.4%	Score 8423;	DB 4;	Length 1665;
Best Local Similarity	99.9%	Pred. NO. 0;		
Matches 1612;	Conservative	0;	Mismatches	1; Indels 0; Gaps 0;
Qy	6356	PPSMQVTTIEDVQAQTGGTAQFAEIIIEGDPQPSVTWKDSVQLVDSTRLSQQOEGTYSIV	6415	
Db	53	PPSMQVTTIEDVQAQTGGTAQFAEIIIEGDPQPSVTWKDSVQLVDSTRLSQQOEGTYSIV	112	
Qy	6416	LRHVASKDAGVYVTCIAQNTGQVLCCKAELLVGGNPEPDKQSHRRKLSHYEVEKEIG	6475	
Db	113	LRHVASKDAGVYVTCIAQNTGQVLCCKAELLVGGNPEPDKQSHRRKLSHYEVEKEIG	172	
Qy	6476	RGVFGFVKRVQHGKNIKILCAAKFIPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETRK	6535	
Db	173	RGVFGFVKRVQHGKNIKILCAAKFIPLRSRTAQAYRERDILAAALSHPLVTGLLDQFETRK	232	
Qy	6536	TLILILECSSEELLDRLYRGVGVTEAEVKYIQQLVEGLHSHGVHLHDIKPSNIIUM	6595	
Db	233	TLILILECSSEELLDRLYRGVGVTEAEVKYIQQLVEGLHSHGVHLHDIKPSNIIUM	292	
Qy	6596	VHPAREDIKICDFGPAQNITPAELQFSQYSGPEFVSPIIQONPVSEASDIWAMGVISYL	6655	
Db	293	VHPAREDIKICDFGPAQNITPAELQFSQYSGPEFVSPIIQONPVSEASDIWAMGVISYL	352	
Qy	6656	SLTCSPPAGESDRATLLNVLEGRVSWSSPMAAHLSEDAKDFIKATLQAPQAPSAQC	6715	
Db	353	SLTCSPPAGESDRATLLNVLEGRVSWSSPMAAHLSEDAKDFIKATLQAPQAPSAQC	412	

QY 6956 LAPGSHSLHSDSPSTRSSACGEAQRLLPAPSGGAPIRDMGHPPQSKQLPSTGGHPG 7015
DB 653 LAPGSHSLHSDSPSTRSSACGEAQRLLPAPSGGAPIRDMGHPPQSKQLPSTGGHPG 712
QY 7016 TAQPERPSDPSWGQAPAPCHPKQSAPOEGCSPPHAPACPPGPPGSCKEAPLVSS 7075
DB 713 TAQPERPSDPSWGQAPAPCHPKQSAPOEGCSPPHAPACPPGPPGSCKEAPLVSS 772
QY 7076 PFLGQAPAPAPAKASPPILDSKMGCDISLPRKPGPCSSPGSASQASSSSQVSLRVGS 7135
DB 773 PFLGQAPAPAPAKASPPILDSKMGCDISLPRKPGPCSSPGSASQASSSSQVSLRVGS 832
QY 7136 SQVTEPGPSLDAEGTQBAEDLSSTPTLQRPQQAATWRKPSLGGRGYAGVAGTGA 7195
DB 833 SQVTEPGPSLDAEGTQBAEDLSSTPTLQRPQQAATWRKPSLGGRGYAGVAGTGA 892
QY 7196 FGGDAGGMLGQGPMMARITAWVSQSEEEQBARAESQEEQBARAESPLPOVSARVP 7255
DB 893 FGGDAGGMLGQGPMMARITAWVSQSEEEQBARAESQEEQBARAESPLPOVSARVP 952
QY 7256 EYGRAPTRSSPEPTPWEDIGVSLVQIRDLSDGDAADTISLDISEVDPAYLNLSDLYDI 7315
DB 953 EYGRAPTRSSPEPTPWEDIGVSLVQIRDLSDGDAADTISLDISEVDPAYLNLSDLYDI 1012
QY 7316 KYLPFEFMI FRKVPKSAQPEPPSPMAEBELAFPEPTWPWPGELGPHAGLEITERSEVD 7375
DB 1013 KYLPFEFMI FRKVPKSAQPEPPSPMAEBELAFPEPTWPWPGELGPHAGLEITERSEVD 1072
QY 7376 ALLAAAVGRKKWSSPSRSLFHPGRHLPLDEPAELGLRERVKASVEHISRLKGRPEG 7435
DB 1073 ALLAAAVGRKKWSSPSRSLFHPGRHLPLDEPAELGLRERVKASVEHISRLKGRPEG 1132
QY 7436 LEKEGPPRRKPGGLASFRSLGSLKSWDRAPTFRLSDSETVVLGQSVTLACQVSAQAQAT 7495
DB 1133 LEKEGPPRRKPGGLASFRSLGSLKSWDRAPTFRLSDSETVVLGQSVTLACQVSAQAQAT 1192
QY 7496 WSKGAPLESRRVLISATLKNFQLTILVVVAEDLGVTCSVNALGTVTGVLRAE 7555
DB 1193 WSKGAPLESRRVLISATLKNFQLTILVVVAEDLGVTCSVNALGTVTGVLRAE 1252
QY 7556 RPSSSPCPDIGVYADGVLVWKPVESGPTVIYVQCSLEGSSWTTLASDIFDCCLTSK 7615
DB 1253 RPSSSPCPDIGVYADGVLVWKPVESGPTVIYVQCSLEGSSWTTLASDIFDCCLTSK 1312
QY 7616 LSRGCTYTFRTACVSKAGMPYSSPSEQVLLGGPSHLASEESQGRSAQPLPSTKTFAPQ 7675
DB 1313 LSRGCTYTFRTACVSKAGMPYSSPSEQVLLGGPSHLASEESQGRSAQPLPSTKTFAPQ 1372
QY 7676 TQIQGRFSVVRQCWEKASGRALAAKIIIPYHPKDKTAVLREYALKGLRHPLHAQLHAAY 7735
DB 1373 TQIQGRFSVVRQCWEKASGRALAAKIIIPYHPKDKTAVLREYALKGLRHPLHAQLHAAY 1432
QY 7736 LSPRLHLILELCSGPELLPCLABRASYESRVKDYLMQMLSATOYLHNOHLLHDLRSE 7795
DB 1433 LSPRLHLILELCSGPELLPCLABRASYESRVKDYLMQMLSATOYLHNOHLLHDLRSE 1492
QY 7796 NMIIITYNLLKVVDLGNQSLQSEKVLPSDKFKOYLETMAPELLEGQAVPQTDIWAIGV 7855
DB 1493 NMIIITYNLLKVVDLGNQSLQSEKVLPSDKFKOYLETMAPELLEGQAVPQTDIWAIGV 1552
QY 7856 TAFIMLSAEPVSSSGARDLQRLKGLVRLSRCYAGLSGGAVAFRLSTLCAQPWGRPCA 7915
DB 1553 TAFIMLSAEPVSSSGARDLQRLKGLVRLSRCYAGLSGGAVAFRLSTLCAQPWGRPCA 1612
QY 7916 SSCLOCPWLTEGPACSPAPVTFTFARLRFVVRNKRKRALLYKRHNLAQVR 7968
DB 1613 SSCLOCPWLTEGPACSPAPVTFTFARLRFVVRNKRKRALLYKRHNLAQVR 1665

RESULT 4

US-09-858-664A-3

; Sequence 3, Application US/09858664A

; Patent No. 6482624

; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO00927-CIP
; CURRENT APPLICATION NUMBER: US/09/858,664A
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: 09/711,134
; FILING DATE: 2000-11-11
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 846
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-858-664A-3

Query Match 2.8%; Score 1173; DB 4; Length 846;
Best Local Similarity 34.9%; Pred. No. 1.4e-54;
Matches 313; Conservative 121; Mismatches 316; Indels 146; Gaps 29;

QY 6357 PSMQVTTLEDVQAQGTGAQPEAIIIEGDPQPSVWYKDSVQLVDSVTRLSQQOEGTYSVLV 6416
DB 1 PRFESIMEDVEVGAGETARFAVVVEGKPLPDIMMYKDEVLLTESHSVSVYVEECSLVV 60
QY 6417 RHVASKAGVVTCLAQNTGGQVLCKAELLVLGGDN----EPDSEKQSHR-RKLHSFYEVK 6471
DB 61 LSTGAQGGVYTCTAQNLAGESVCAELAVHSAQTAMEVEGVGEDEHGRRLSDFYDIH 120
QY 6472 BEIGRGVGFVVRVQHKGNKILCAAKTIPURSTRQAQYRERDILAAHSLPVLTLGLDQF 6531
DB 121 QEIGRGAFSVLRRIVERSGGLFAAKTIPSOAKPKASARREARLLARLQHDCLVLYFHEAF 180
QY 6532 ETRKTLILLELCSSEELLDRLYEKGVVTEAEVKVYLQOLVEGLHYLHSHVTLHLDIKPS 6591
DB 181 ERRGLVITELC-TESELEKIAKPTVCESEIRAYMRQVLEGLHYLHSHVTLHLDIKPS 239
QY 6592 NILMVHPA--REDIKICDFGFAQNIPTAELOFQSOYGSPEFVSPEIIOQNVPVSEASDIWAM 6649
DB 240 NLLVWDGAGCEQVRIICDFGNAQELTPEQYCYQGTPEFVAPFVIVNQSPVSGVTDIWPV 299
QY 6650 GVISYLSLTCSPPAGSDRATLNLVLEGRVSWSPMAAHLSEDAKDP- IKATLQRAPOA 6708
DB 300 GWAFLCLTGISPPVGENDRITLMNIRNVAFBETTFLSLSREARGFLIKVLVQ--DRL 357
QY 6709 RPSAAQCLSHPWFLKMPABEAHPINTKQLFLIARSRWORSLSYKSLVMRSIPELLR 6768
DB 358 RPTAEETLEHPWFKTQAKGAE---VSTDHLKFLSRRRWORSQISYKCHLVIRIPELLR 414
QY 6769 GPPDSPSLGVARHLCDRTGGSSSSSSSSSDNEL-----APPARAK-SLPPSPVTH 6816
DB 415 APPERVVMTPRR--PPPSGGLSSSSDSEEELEELPSVPRPLQPEFGSRVSLTDIPTED 473
QY 6817 SPLPHRGFLRPSASLPEEABASRSTETAPAPASPRGAGPPAAQCGCVPRHSVTR----- 6871
DB 474 EALGTPTGAATPMDWQEQGRAPSDQOEAPSPALPSGOEPAA-GASPRRGEELRGSSA 532
QY 6872 -----SLFYHQAG--ESPEHGALAPG-----SRRHPPARRRHLLK 6903
DB 533 ESALPRAGPRELGRGLHKAASVELPQRSPGPGATRLARGGLGEVEAQRLOALQRLLR 592
QY 6904 GGYTAGALPGLREPLMEH-----RVLEEEAARHEEQATL-----LAKAPSFETALR 6948
DB 593 GGPEDGVSGLRGFLLESLSGGRARDPRMAAASSEAAHPHPPLLENRGLQKSSFSQGEA 652
QY 6949 LPASGTHLAPGCHSHSLHSDSPSTR-----PSSACGEAQRLLPSPASGAPIRDMGHPPG 7004
DB 653 EP-RGRHRRAGAPLEIPVARLGARRLOESFSLSALSEAQ--PSSPA-----RPSAP 700
QY 7005 KQLPSTGGHPCTAQPERPSPDPSWGQAPFCHPKQSGAPQOEGCSPPHAPACPPGSPPPG 7064
DB 701 X--PST---PKSAEPSATTESDAPQPPAP--QPAQDKAPRPPRPPVRSKAPAP---PPQ 749


```
Db 240 NLLVWDGAGGQQVRIICDFGNAQELTGPCEQYQYGTPEFVAPFVQSVGVTDIWVP 299
Qy 6650 GVISYLSITCSPSPAGSDRATLNVLEGRVSWSSPMAHLSSEDAKF- IKATLQAPQA 6708
Db 300 GVWAFCLTIGISPFVGENDRITLNRINYNVAFETTFPLSLSRARGFLIKVLVQ--DRL 357
Qy 6709 RPSAAQCLSHPWFLKSPAEAAHINTKQLAFLARSRWORSLSYKSLVMRSIPELLR 6768
Db 358 RPTAETLEHFWFKTOAGAB---VSTDHLKLFSLRRWQRSQISYKCHLVIRIPELLR 414
Qy 6769 GPPDSPSLGVARHLCDRTDGGSSSSSSSDNEL-----APFARAK-SLPPSPVTH 6816
Db 415 APPERVVMTFRR-PPPSGGLSSSDSEEELEELPSVPRLPQPEFGSRVSLTDIPTED 473
Qy 6817 SPLHLPRGFLRPSASLPREABASRSTAPAPPASPEGAGPAAQCGVPRHSVIR----- 6871
Db 474 EALGTPETGAATPMDWQGRAPSDQOQAPSPALPSGQFPAA-GASPRRGELRRGSSA 532
Qy 6872 -----SLFYHQAG-ESPEHGALAPG-----SRRHPARRRHLK 6903
Db 533 ESALPRAGPRGLRGLHKAASVELPQRSFGATRLARGGLGEVQAORLQALRQLLR 592
Qy 6904 GGYIAGALPGLREPLMEH-----RVLEEEAAREEOATL-----LAKAPSFETALR 6948
Db 593 GGPEDGKVSGLRGLPGLSLGGRARDPRMARAAASSEAPHQPLENRLGLKSSFSQGEA 652
Qy 6949 LPASGTHLAPGHSHLSHDSFSTR-----PSSEACGEAORLPASPGAPIRDWGHQGS 7004
Db 653 EP-RGRHRRAGAPLEIPVARLGARRLQESPLSALSSEAQ--PSSPA-----RPSAP 700
Qy 7005 KQLPSTGCHGCTAOPERPSPDSPWGQAPFCHPKQGSAPQBGCPHPHAPVAPCPGSPFPG 7064
Db 701 K--PST---PKSABFSATTPSDAQPAP--QPAQDKAPEPRPREVRASKAP-----PPQ 749
Qy 7065 SCK--EAPLVFSPSFLGQ-----POAPPAPAK-----ASPLDLSKMP 7100
Db 750 ALQTLALPLTFYAQIIQSLQSLGHAQSPSQAPPSBPKEPAAVFAVASFPP----- 802
Qy 7101 GDLSLPGPKPGCPSPGASQASQSSQVSSLRVSSQVGTGPGSLDAGHTQAE 7156
Db 803 -----PGAPEKRVFSAGGPPVLAEKARVPT-----VPPRPGSSLSSENLESE 846

RESULT 7
US-10-006-011A-2
; Sequence 2, Application US/10006011A
; Patent No. 6821947
; GENERAL INFORMATION:
; APPLICANT: Iozzo, Renato V.
; TITLE OF INVENTION: Endorepellin: methods and compositions
; TITLE OF INVENTION: for inhibiting angiogenesis
; FILE REFERENCE: 8321-95
; CURRENT APPLICATION NUMBER: US/10/006,011A
; CURRENT FILING DATE: 2001-12-04
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 4391
; TYPE: PRT
; ORGANISM: human
US-10-006-011A-2

Query Match 2.6%; Score 1074; DB 4; Length 4391;
Best Local Similarity 24.2%; Pred. No. 3.1e-48;
Matches 653; Conservative 273; Mismatches 947; Indels 830; Gaps 140;

Qy 32 QIVGNPTQVSWKXQDPVTAGARPLAQ-----DGLVRLTILDLALGDSQY 80
Db 1436 QITGNMILVA-----SQALQGPERSYEIMREFEWFRRDPQPATREHLLMALLDLDEL 1491
Qy 81 VCRARNATGEAPAAVGLVDAAEACBQA-PHFLLRPTSIIRVREGSEATFRCR-----VG 134
```

```
Db 1492 LIRA-----TFPSVPLVASISAVLSVAQPGPSNRPRALAEVE-----CRCPPGVIG 1538
Qy 135 GSPR---PAYSWSKDGRRLG-----EPDG-----PRVRVEELCEASALRIRAAAPRDGG 180
Db 1539 LSQDCAPGVTRGSLYLGHCELCENGHSDLCHP-----ETGACSQCOHNA-----GE 1589
Qy 181 TYEVAENPLGAASAAALVVDSDAATASRPGTSTAALLAHLQRRREARMAEG---APA 237
Db 1590 FCCLCAPYYGDATAGT---PEDCQPCACPLTNPENMF---RTCESLACAGYRCTAC 1641
Qy 238 SPSTG-----TFTCTVTEGKHARLSYV 261
Db 1642 EPGYTGQYQCQCGPYGYNPVSQGGCLPETNQAPLVVEVHPAKSIYPVQGSHS-LRCQV 1700
Qy 262 TGEPKPTVW-KKDGQLVTEG--RRHVVEDAONFVLKILFCQSDRGLTYCTASNL-V 317
Db 1701 SGSPPHYFWSREDGRVPVSGTQQRH-----QGSSELHFPSPQPSDAGVICTCRNLHQ 1753
Qy 318 GQTVSSVLVVRBPVAPFKRLQD---LEVREKESATFLC--EVPQPSSTEAMFKEETRL 372
Db 1754 SNTSRAELLVTEAPSKPITVVEEQRSQVRPGADVTFICTAKSKSPAYTLVW---TRL 1809
Qy 373 WASAKYIEEGTERRLTVRNVSDADDVAVICETPEGSRTVAELAVOGLNLLRKLPRKA- 431
Db 1810 -HNGKLPTRAMDFNGILTIRNVQLSDAGTYVC-----TGSNFMFAMDQGTAT 1854
Qy 432 --VRVGTAMFCVELAVPVGPHVWLRNQEVEVAGGRVAISAEGRTHLTITISQCCLDEVGQ 489
Db 1855 LHVOASGT-----LSAPVVSIIH-----PPQITVQP-----GQ 1881
Qy 490 VAFMAGDCQSTTRFCVSAPRKPLPPVDPVVKARMESSVLWSPPHGRP--VTIDG 547
Db 1882 LA-----EPKRSATGSP-----TPTLEWTGGPGGOLPAKQAIHG 1915
Qy 548 YLVEKKLGTYYTWRICHEAENWATPELTVDVAEENGFQFVRSALNSFGQSPYLEFFCTV 607
Db 1916 GILR-----LPAVEPTDQA-----QYLCRAHSSAGQO----- 1942
Qy 608 HLAPKLAVRTPLKAVQAVEGG---EVTFSVDLTIVASAGFWFLDQCALKASSVYBIHCRT 664
Db 1943 -----VARAVLHVHGGGRVQVSPERT-----QVHAGRT 1972
Qy 665 RHTLTIREVPASLHGAQLKFVANGIESSIRMEVFAAPGLTANKPPAAAAREVLARLHEEA 724
Db 1973 -----VRLYCRAA----- 1980
Qy 725 QLLAELSDQAAAVTMLKDGRTLSPKPYEQASAGR---RVLLVRDVARDDAGLYECV-- 779
Db 1981 -----GVPSATITWRKEGSLPP-----QARERTDIATLLIPAITADAGFYLCVAT 2028
Qy 780 SRGRIAYQLSVQGLARFLHKDMAGSCVDVAVAGGPAQFECET-----SEAHVHVHWY 831
Db 2029 SPACTAQRMQVVVLSA---SDASPPGVKIBESSPSVTEGTDLNLCVVAGSAHAQVTWY 2085
Qy 832 KQGMEL-GHSGERFLOEDVGRHRLVAATVTRQDEGYTSCV--GEDSVDFRLRVS----- 884
Db 2086 RRGGSLLPHT-----QVHGSRLRL--PQVSPADSGEYVCRVENGSGPKESITVSVLHG 2137
Qy 885 -----EPKVVFAKEQLARRKLQABAGASATLSCRV-AQAQTEVTWY 924
Db 2138 THSGFSYTPVPGSTRPRIEPSSSHVAE-----GQTLNLCVVQQAHAQVTHW 2186
Qy 925 KDGKLSSSSKVCMCEATGCTRLRVVQOAGADAGEYSCAAGORLSFLHDYK---EPKV 981
Db 2187 KRGSLSLPAH---QTHGSLRL--HQVTPADSGEYVCHVVGTSGPLEASVLVTIEASVI 2240
Qy 982 ---FAKQVHAHSEVOAEAGANATLSCEVA-QAQAEMVMYKDGKLSLKVHVEAKGCR 1037
Db 2241 PGPIPPVRIESSSTVAEGQTLDLSCVVAGQHAQVTVYKRG---GSLPARHQVRG--S 2294
Qy 1038 RLVVQOAGKTDAGDYSCEA-RGQVSPRLHIT-----EPKMMFAKE 1077
Db 2295 RLYIFQASPADAGQVVCRAENGMEASITVTVTQGANLAYPAGSTQPIRIEP----- 2347
```



```
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 549
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-697-263-6

Query Match          2.2%; Score 909.5; DB 4; Length 549;
Best Local Similarity 38.3%; Pred. No. 1.1e-40;
Matches 204; Conservative 88; Mismatches 181; Indels 59; Gaps 12;

QY 6274 TGTSEAPVPPRPVLPOLLHEGPEQPEAIARAQEWTPVRMEGAAMPAGGT-----GE 6326
Db 25 TVKSSKSPPEPQVLEHGTLE-EAPAMLDKPDIVVVGQ--PASVTVFNHVEAQ 81
QY 6327 LMDVHSHVVRTTQRTTYQAID-----THTAR----- 6355
Db 82 VVMRSRCGALLEARAGVYELSQPDDQYCLRICRVSRDRMGALTCARNRHGTQTCVTL 141
QY 6356 ----PPSMQVTIEDVQAOTGGTAQFEAIIIEGDPQPSVTWYKDSVOLVDSTRLSSQQOEGTT 6411
Db 142 ELAEAPRFESIMEDVEVGAGETARFAVVVEGKPLDIIWYKDEVLLTSSSHVSFVYEENE 201
QY 6412 YSLVLRHVASKDAGVYTCLAQNTGGQVLCKABELLVGGDN----BPDSEKQSHR-RKLHS 6466
Db 202 CSLVLSTGAQDGGVYTCTAQNLAGESVCKAELAVHSAQTAMEVBGVGDEDEHGRRLSD 261
QY 6467 FVEVKEETGRGVGVKRVQVHKGKILCAKPIPLRSRTRAQAYRERDILALSHPLVTG 6526
Db 262 FYDIHQEIGRGAFSLRIVERSSGLEFAAKFIPSOAKPKASARREARLLARLQHDVCVLY 321
QY 6527 LLDQFETKTLILLELCSSELLDLRYKGVVTEAEVKVYIQQLAVEGLHYLHSHGVHLH 6586
Db 322 FHEAFERRGLVITELC-TEELLERIAKPTVCESEIRAYMRQVLEGHILHSHGVHLH 380
QY 6587 DIKPSNILMVHPA--REDIKICDFGAQNTPAELQFSGYSGSPFVSPSEIIQQNPVSRAS 6644
Db 381 DVKPNLLAVDGAAGEQQVRCDFGNAQBELTPGEQYCYQYGTPEFVAPEIVNQSPVSGVT 440
QY 6645 DIWAMGVLSYLSLTCSPPAGESDRATLLNVLEGRVSWSSPMAHLSDAKDF-IKATLY 6703
Db 441 DIWPGVVAFLCLTGISFPVGENDRITLLMIRNYNAVEETFTLSLSREARGFLIKVLVQ 500
QY 6704 RAPQARPSAAOCLSHPWFLKSPAEAEHFINTKOLKFLARSRWORSLSMYK 6755
Db 501 --DLRPTAEETLEHPWFKTQAKGAE---VSTDHLKFLSRRRWQRSISYK 547
```

```
RESULT 12
US-08-826-267-2
; Sequence 2, Application US/08826267
; Patent No. 5994070
; GENERAL INFORMATION:
; APPLICANT: Streuli, Michel
; TITLE OF INVENTION: No. 5994070el TRIO Molecules and Uses Related Thereto
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/826.267
; FILING DATE: 1997
; CLASSIFICATION: 800
```

```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/014.214
; FILING DATE: 27 MARCH (1996)
; ATTORNEY/AGENT INFORMATION:
; NAME: Amy E. Mandragouras
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: DFN-010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2860 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-826-267-2

Query Match          2.2%; Score 892.5; DB 2; Length 2860;
Best Local Similarity 24.7%; Pred. No. 9.3e-39;
Matches 318; Conservative 209; Mismatches 493; Indels 265; Gaps 42;

QY 5562 DSDSDSTKPSASPRHGRSR---PSSIQSSSSSESD--GDARGEIPIYVVTADYPL- 5614
Db 1758 DSDSDAATPODETVEERNEGLSGTLSSSSSQMCGEEGE-----EGADAVPLP 1811
QY 5615 ---CAEODAITLRGEYVEVLDAAHPLRLWLVRTKTKS-SPSRQGWYSP-----AY 5661
Db 1812 PPMAIQHSLLQPOSQ-----DDKASSRLLR--PTSSETPSAAELVSAIEELVKSXAL 1864
QY 5662 LDRRLKLSPEWGAAEPEF-PGE-----AVSEDEYKARLSS-----VTQELLSSEQA 5707
Db 1865 EDRPSSLVDOQSSSPFNPSDNSLLSSSPIDEMEERKSSSLKRRHYVLQELVETRD 1924
QY 5708 FVEELQFLQSHHLOHLERCPHVPIAVAGQXAVIPRNVRDIGRFHSS-FLOELQOCDTD-D 5765
Db 1925 YVRDLGVVWVGMA-LMKEDGVDDMKGDKDQIVFGNTHQIVDHRDFFLGELEKLEDPE 1983
QY 5766 DVAMCFIKNOAAPEQYLEFLVGRVQAESVVVSTAIQEFYKKAEEALLAGDPSPPPPPL 5825
Db 1984 KLGSLFKVHERRLHMYIAYCONKPKSEH-IVSEYIDTFFEDLKQR-----L 2028
QY 5826 QHYLE-----QPVERVQRYQALLKELIRNKARNRQNCALLEQAYAVVSALPORAENKLH 5879
Db 2029 GHRQLTDLLIKPVQIRMKYQLLKDFLYSKKASLDTSELERAVEVCMVPRCNDMMN 2088
QY 5880 VSLMENYPTLEALGEPFIRQGHFIWEGAPGARMPWGHNRHVLFRNHLVICKPRDRSR 5939
Db 2089 VGRLQGPDKIVAQGKLLQDTFLVTDQDAGL-LP-RCRERRIFLFSQIVIFSEPLDKKK 2146
QY 5940 TDTV-SYVFRNMKLISSIDLNDQVGDRAFEVWQEREDSVRKYLLOARTAIKSSWVKE 5998
Db 2147 GFSMPGFLFRNSIKVSCLCLEENVEDPCFKALTSTRTGDDVVETFILSSSPSVQTTWIHE 2206
QY 5999 ICGI---QORLALFVWRPDPFEEL-----ADCTAELGETVKLACRVTTGTPKPKVISW 6047
Db 2207 INQILENQRFNALTSFIEYQNRNHSGGGGGGGGAAGVG-----AAAAAGPPVAAAAT 2261
QY 6048 YKDGKAVQVDPHILIEDPDGSCALILDSLTGVDGGYMCFAASAAGNCSTGLKILVQVP 6107
Db 2262 VAAPAAAAAPP-----ARAGAG-----P 2279
QY 6108 PRFVNKVRASPFVEGEDAQFTCTIEGAPYQIRWYKDGALLTGNKFTQLSEPRSGLLVL 6167
Db 2280 P-----GSFSL--SDTTPCWSPLQPRARQR-----QTRQSSSSSSSNISTM 2320
QY 6168 VI-----RAASKEDGLGYECELVNRGLSARASAEIRIQSPMLQAEQEC----- 6210
Db 2321 LVTHDYTAVKDEDEINVYQGEVVQILASNQNMELVFPRA---ATDQCPAAAGWIPGVLG 2376
QY 6211 HRQLVNAVEDTTLERADQEVTSVLKRLGPK--APGPFSTGDLTGPCPRGAPALBETG 6268
Db 2377 HTSAVIVENPDGTLKKSTSWHTALRLRKSEKKDKGREGKL-----ENG 2422
```



```
QY 6269 SPPVTGTSAPVPRVPDPLLHSGPEQPEPAIARAQEWTPVIRMEGANWPGAGTGELL 6328
Db 2423 YRKSEGLSN-----KVSVKLLNP-----2441
QY 6329 WDVHSHVRETTRQTYTQAIDHTARPPSMOVTIEDVQAOTGGTAQPEAIEGDPQPSV 6388
Db 2442 -----NYIDV-----PPEFVIPLEVTCTGETVVLRCVCGRPKASI 2480
QY 6389 TWYKDSVOLVD-----STRLSQQQGGTYSVLVLRHVASKDAGVYCTLAQNTGGQVLCKAE 6443
Db 2481 TWKGPEHNTLNNDGHYSISYSLGGEAT---LKIIVGVTTEDDGIYTCIAVNDMGSSASSAS 2537
QY 6444 LLVLGGDNEPSEKSHRKLHSFYEVKEEIGRGVFGFKVRVQHKGKILCAKFIPLRS 6503
Db 2538 LRVLFGP---MDGIMVTWKDNFDSFYSEVAELGRGRFSVVKCDQGTGRAVATKFNKLL 2595
QY 6504 RTRAQYRERDILAALSHPLVTGLDQPETRKTLILILELSCSSEBLLDRLYRKGVVTEAE 6563
Db 2596 MKRDQVTHELGILQSLQHPLLVGLLDTFTPTSYILVLEMAQGRLLDCVVRWGLTEGK 2655
QY 6564 VKVYIQQLVEGLYHSHGVHLDTKPSNLMVHP-AREDIKIDFGFAQNTTPAELQFS 6622
Db 2656 ITRAHLGEVLEAVRYLHNCRIAHLDLKPENILVDESIAKPTIKLADFGDAVQLNTTYIHQ 2715
QY 6623 QYSGSPFVSPELQONPVSEASDIWAMGVISYLSLTCSSPPAGESDRATLNLVLEGRVSW 6682
Db 2716 LLGNPEFAAPEILGNPVSLSDTWSVGLTYVLLSGVSPFLDDSVETCLNICRLDFSP 2775
QY 6683 SSPMAAHLSEDAKOFIKATLQAPQAPSAOCLSHPMFLKSMPEAEAHFINTKOLKELL 6742
Db 2776 PDDYFKGVSKAKFVFCVLLQEDPAKRSALALQEQW-LQAGNGRSTGVLDTSRLTSFI 2834
QY 6743 ARSRMQ---RSLMSYKSLVMRSIP 6764
Db 2835 ERKHQNDVRPIRSIKNFLQSRLLP 2859
```

RESULT 13

```
US-09-858-664A-13
; Sequence 13, Application US/09858664A
; Patent No. 6482624
```

```
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui, et al.
```

```
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000927-CIP
```

```
; CURRENT APPLICATION NUMBER: US/09/858,664A
```

```
; PRIOR FILING DATE: 2001-05-17
```

```
; PRIOR APPLICATION NUMBER: 09/711,134
```

```
; NUMBER OF SEQ ID NOS: 33
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 13
```

```
; LENGTH: 414
```

```
; TYPE: PRT
```

```
; ORGANISM: Homo sapiens
```

```
US-09-858-664A-13
```

```
Query Match 1.4%; Score 592.5; DB 4; Length 414;
```

```
Best Local Similarity 34.5%; Pred. No. 7.6e-24;
```

```
Matches 145; Conservative 69; Mismatches 191; Indels 15; Gaps 6;
```

```
QY 6356 PPSMQVTIEDVQAOTGGTAQPEAIEGDPQPSVTWKDSVOLVD-----STRLSQQQGGT 6410
Db 1 PPEFVIPLEVTCTGETVVLRCVCGRPKASITWKGEHNTLNNDGHYSISYSLGGEAT 60
```

```
QY 6411 TYSVLVLRHVASKDAGVYCTLAQNTGGQVLCVCAELVLLGGDNEPSEKSHRKLHSFYEV 6470
```

```
Db 61 ---LKIIVGVTTEDDGIYTCIAVNDMGSSASSASLRVLGPG--MDGIMVTWKDNFDSFYSE 115
```

```
QY 6471 KEETIGRGVFGFKVRVQHKGKILCAKFIPLRSRTRAQYRERDILAALSHPLVTGLDQ 6530
```

```
Db 116 VAEELGRGRFSVVKCDQGTGRAVATKFNKLMKRDQVTHELGILQSLQHPLLVGLLDT 175
QY 6531 PETRKTLLILILELSCSSEBLLDRLYRKGVVTEAEVKVYIQQLVEGLYHSHGVHLDTKP 6590
Db 176 FETPTSYILVLEMAQGRLLDCVVRWGLTEGKIRAHLEGEVLEAVRYLHNCRIAHLDLKP 235
QY 6591 SNILMVHP-AREDIKIDFGFAQNTTPAELQFSQYSGSPFVSPELQONPVSEASDIWAM 6649
Db 236 ENILVDESIAKPTIKLADFGDAVQLNTTYIHHQLLGNPEFAAPEILGNPVSLSDTWSV 295
QY 6650 GVISYLSLTCSSPPAGESDRATLNLVLEGRVSWSSPMAAHLSEDAKOFIKATLQAPQAR 6709
Db 236 GVLITVLLSGVSPFLDDSVETCLNICRLDFSPDDYFKGVSKAKFVFCVLLQEDPAKRS 355
QY 6710 PSAAOCLSHPMFLKSMPEAEAHFINTKOLKELLARSRWQ---RSLMSYKSLVMRSIPBL 6766
Db 356 PSALALQEQW-LQAGNGRSTGVLDTSRLTSFIERRKHQNDVRPIRSIKNFLQSRLLP 414
```

RESULT 14

```
US-10-274-978-14
```

```
; Sequence 14, Application US/10274978
```

```
; Patent No. 6670164
```

```
; GENERAL INFORMATION:
```

```
; APPLICANT: WEI, Ming-Hui, et al
```

```
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
```

```
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
```

```
; TITLE OF INVENTION: THEREOF
```

```
; FILE REFERENCE: CL000927-CIP-DIV
```

```
; CURRENT APPLICATION NUMBER: US/10/274,978
```

```
; PRIOR FILING DATE: 2002-10-22
```

```
; PRIOR APPLICATION NUMBER: 09/858,664
```

```
; PRIOR FILING DATE: 2001-05-17
```

```
; PRIOR APPLICATION NUMBER: 09/711,134
```

```
; PRIOR FILING DATE: 2000-11-14
```

```
; NUMBER OF SEQ ID NOS: 34
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 14
```

```
; LENGTH: 414
```

```
; TYPE: PRT
```

```
; ORGANISM: Human
```

```
US-10-274-978-14
```

```
Query Match 1.4%; Score 592.5; DB 4; Length 414;
```

```
Best Local Similarity 34.5%; Pred. No. 7.6e-24;
```

```
Matches 145; Conservative 69; Mismatches 191; Indels 15; Gaps 6;
```

```
QY 6356 PPSMQVTIEDVQAOTGGTAQPEAIEGDPQPSVTWKDSVOLVD-----STRLSQQQGGT 6410
Db 1 PPEFVIPLEVTCTGETVVLRCVCGRPKASITWKGEHNTLNNDGHYSISYSLGGEAT 60
```

```
QY 6411 TYSVLVLRHVASKDAGVYCTLAQNTGGQVLCVCAELVLLGGDNEPSEKSHRKLHSFYEV 6470
```

```
Db 61 ---LKIIVGVTTEDDGIYTCIAVNDMGSSASSASLRVLGPG--MDGIMVTWKDNFDSFYSE 115
```

```
QY 6471 KEETIGRGVFGFKVRVQHKGKILCAKFIPLRSRTRAQYRERDILAALSHPLVTGLDQ 6530
Db 116 VAEELGRGRFSVVKCDQGTGRAVATKFNKLMKRDQVTHELGILQSLQHPLLVGLLDT 175
```

```
QY 6531 PETRKTLLILILELSCSSEBLLDRLYRKGVVTEAEVKVYIQQLVEGLYHSHGVHLDTKP 6590
```

```
Db 176 FETPTSYILVLEMAQGRLLDCVVRWGLTEGKIRAHLEGEVLEAVRYLHNCRIAHLDLKP 235
```

```
QY 6591 SNILMVHP-AREDIKIDFGFAQNTTPAELQFSQYSGSPFVSPELQONPVSEASDIWAM 6649
Db 236 ENILVDESIAKPTIKLADFGDAVQLNTTYIHHQLLGNPEFAAPEILGNPVSLSDTWSV 295
```

```
QY 6650 GVISYLSLTCSSPPAGESDRATLNLVLEGRVSWSSPMAAHLSEDAKOFIKATLQAPQAR 6709
```

```
Db 296 GVLITVLLSGVSPFLDDSVETCLNICRLDFSPDDYFKGVSKAKFVFCVLLQEDPAKRS 355
```

```
QY 6710 PSAAOCLSHPMFLKSMPEAEAHFINTKOLKELLARSRWQ---RSLMSYKSLVMRSIPBL 6766
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 17:25:27 ; Search time 575 Seconds
(without alignments)
5359.490 Million cell updates/sec

Title: US-10-077-130-5

Perfect score: 41273

Sequence: 1 MDQPFSCAPFLTRPKAFV.....RNREKRALLYKRNHLAQVR 7968

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

1: geneseqp1980e:*

2: geneseqp1990e:*

3: geneseqp2000e:*

4: geneseqp2001s:*

5: geneseqp2002s:*

6: geneseqp2003as:*

7: geneseqp2003bs:*

8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	41273	100.0	7968	6	ABG76187 Human ser
2	32134	77.9	6620	7	ADJ70485 Human hea
3	13710	33.2	2630	6	ABG76186 Human ser
4	13528	32.8	2596	4	AAB30569 A splice
5	10519.5	25.5	4691	5	ABP70084 Human NOV
6	10506	25.5	4675	5	ABP70085 Human NOV
7	9834.5	23.8	2328	6	ABP58227 Human cel
8	8423	20.4	1665	5	AAO15372 Human myo
9	8410	20.4	1665	5	AAE24151 Human kin
10	8407	20.4	1618	4	AAB85504 Human pro
11	8403	20.4	1610	4	AAB30568 A full le
12	7073	17.1	1351	4	AAB30567 Amino aci
13	6619.5	16.0	1596	7	ADJ69186 Human hea
14	4533	11.0	871	4	AAB30570 A splice
15	4529	11.0	871	5	AAE16274 Human kin
16	3878.5	9.4	31267	6	ABG74786 Human RGS
17	3580.5	8.7	26926	6	AAU05396 Human tit
18	3577.5	8.7	26926	8	ADQ17316 Human sof
19	3427	8.3	34350	8	ADQ89964 Antagonist
20	2489	6.0	6642	8	ADN22360 Bacteriol
21	2447.5	5.9	548	4	AAB30571 A full le
22	2415.5	5.9	548	4	AAB30572 A full le
23	2360	5.7	4412	3	AAV53666 Sequence
24	2328.5	5.6	3267	7	ADQ99075 Human KPP
25	2324	5.6	3208	7	ADG47676 Human NOV

26	2324	5.6	3208	8	ADJ78946	ADJ78946	Human NOV
27	2319	5.6	3268	7	ABD79959	ABD79959	Human kin
28	2270	5.5	3252	7	ADG47674	ADG47674	Human NOV
29	2270	5.5	3252	8	ADJ78944	ADJ78944	Human NOV
30	2203	5.3	3262	7	ABD79961	ABD79961	Mouse ser
31	2192.5	5.3	3262	8	ADR66062	ADR66062	Human pro
32	2192.5	5.3	5614	8	ADR66404	ADR66404	Human pro
33	2135	5.2	2380	5	AAE19160	AAE19160	Human kin
34	2127	5.2	2286	4	AAB65635	AAB65635	Novel pro
35	2127	5.2	2286	8	ADJ29242	ADJ29242	Human MAR
36	2102	5.1	5635	5	ABP60991	ABP60991	Novel hum
37	2085.5	5.1	3186	7	ADG47672	ADG47672	Human NOV
38	2085.5	5.1	3186	8	ADJ78942	ADJ78942	Human NOV
39	2085	5.1	416	7	ADM05704	ADM05704	Human pro
40	2072.5	5.0	5636	7	ADJ70089	ADJ70089	Human hea
41	2072.5	5.0	5636	7	ADJ81337	ADJ81337	Human hem
42	2072.5	5.0	5636	8	ADK60205	ADK60205	Angiogene
43	2072.5	5.0	5636	8	ADK60506	ADK60506	Angiogene
44	2072.5	5.0	5636	8	ADP73129	ADP73129	Angiogene
45	2067	5.0	2231	7	ABD79962	ABD79962	Human KIA

ALIGNMENTS

RESULT 1

ABG76187
ID ABG76187 standard; protein; 7968 AA.

XX AC ABG76187;

XX DT 09-MAY-2003 (first entry)

XX DE Human serine/threonine or protein kinase 12599.

XX KW Human; enzyme; serine/threonine kinase; protein kinase; 12599;
XX KW cardiovascular disease; heart failure; myocardial infarction;
XX KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma; immunogen;
XX KW blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
XX KW haemolytic anaemia; cellular proliferative disorder; cancer;
XX KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
XX KW psoriasis; inflammatory bowel disease; rheumatoid arthritis;
XX KW multiple sclerosis.

XX OS Homo sapiens.

XX PN US2002168742-A1.

XX PD 14-NOV-2002.

XX PF 15-FEB-2002; 2002US-00077130.

XX PR 15-FEB-2001; 2001US-0269201P.

XX PA (MILL-) MILLENNIUM PHARM INC.

XX PI Kapeller-Libermann R, Acton SL;

XX DR WPI; 2003-298729/29.

XX DR N-PSDB; ABX11642.

XX PT Novel isolated human protein kinase, designated 59079 or 12599

XX PT polypeptide, useful as diagnostic and therapeutic agents for preventing

XX PT cardiovascular diseases, proliferative disorders, and protein kinase

XX PS Claim 8; Page 84-104; 119pp; English.

XX The invention relates to an isolated human serine/threonine or protein

XX kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule

XX comprising at least 85% identity to the nucleic acids appearing as

XX ABX11641 and ABX11642 or their complement, a naturally occurring variant

XX of the kinases or their fragments. Also included are a non-human host

cell containing the nucleic acids, an antibody specific for the proteins, identifying a compound which binds to the kinase (by contacting the kinase or a cell expressing the kinase with a test compound and determining whether the kinase binds to the test compound) and modulating the activity of kinase using the identified compound. The kinases and their encoding nucleic acids are useful as diagnostic and therapeutic agents for preventing a disease or condition associated with an aberrant or unwanted 59079 or 12599 activity in a subject, including cardiovascular diseases such as heart failure, and myocardial infarction; disorders involving blood vessels such as atherosclerosis, and Kaposi's sarcoma; blood platelets disorder such as thrombocytopaenia, and Leukemia; Hodgkin's disease, hemolytic anaemia; cellular proliferative disorders such as cancer; and protein kinase disorders such as autoimmune rheumatoid arthritis, psoriasis, inflammatory bowel disease, and disorders are included in the specification). The kinases, their encoding nucleic acids and antibodies are useful in screening assays, detection assays (e.g. forensic biology), and predictive medicine (e.g. diagnostic assays, prognostic assays, and monitoring clinical trials and pharmacogenomics). The kinases and their encoding nucleic acids are useful as query sequences to perform a search against public databases to identify other family members or related sequences. The present sequence represents kinase 12599

XX Sequence 7968 AA;

Query Match	100.0%;	Score 41273;	DB 6;	Length 7968;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 7968;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	MDQFSGAPRLTRPKAFVSVGKDATLSQIVGNPTFQVSWKXQDPVTAGARFLAQ	60
DB	1	MDQFSGAPRLTRPKAFVSVGKDATLSQIVGNPTFQVSWKXQDPVTAGARFLAQ	60
QY	61	DGDLRLTLDLALGDSGYVCRANVAIGEAFAAVGLQVDAEAAQAPHLRPTIR	120
DB	61	DGDLRLTLDLALGDSGYVCRANVAIGEAFAAVGLQVDAEAAQAPHLRPTIR	120
QY	121	VREGSEATFRCRVGSPPAPVSWSKDGRRLGPDGPRVVRVEELGEASALIRARPRDGG	180
DB	121	VREGSEATFRCRVGSPPAPVSWSKDGRRLGPDGPRVVRVEELGEASALIRARPRDGG	180
QY	181	TYEVAENPLGAASAAAALWDSDAADTASRPGTSTAALLAHLQRRREAMEAGAPASPP	240
DB	181	TYEVAENPLGAASAAAALWDSDAADTASRPGTSTAALLAHLQRRREAMEAGAPASPP	240
QY	241	STGTRTCTVTBKGKARLSCTVTGEPKPTVWKDQQLVTEGRRHVYVEDAQNFKILF	300
DB	241	STGTRTCTVTBKGKARLSCTVTGEPKPTVWKDQQLVTEGRRHVYVEDAQNFKILF	300
QY	301	CKQSDRGLYTCTASNLVGQYSSVLVVRPAVPFKKRLQDLVREKESATFLCEVPQPS	360
DB	301	CKQSDRGLYTCTASNLVGQYSSVLVVRPAVPFKKRLQDLVREKESATFLCEVPQPS	360
QY	361	TEAAWFKEETRLWASAKYGIIEEGTERLTVRNVSADDDAVYICETPGSRTVAELAVQG	420
DB	361	TEAAWFKEETRLWASAKYGIIEEGTERLTVRNVSADDDAVYICETPGSRTVAELAVQG	420
QY	421	NLLRLPKRTAVRVGDTAMFCVELAVPGVPHMLRNQEVVAGGRVAISAGTRHTLTIS	480
DB	421	NLLRLPKRTAVRVGDTAMFCVELAVPGVPHMLRNQEVVAGGRVAISAGTRHTLTIS	480
QY	481	QCCLDEVQVAFMAGDCQTSFRFCVSAAPRKLPPDPVPPVKARMESSVILSWSPPHGE	540
DB	481	QCCLDEVQVAFMAGDCQTSFRFCVSAAPRKLPPDPVPPVKARMESSVILSWSPPHGE	540
QY	541	RPVTIDGVLVEKKLGTWTWIRCHAEAWATPELTADVABEGNFPQFRVSALNSFGOSPY	600
DB	541	RPVTIDGVLVEKKLGTWTWIRCHAEAWATPELTADVABEGNFPQFRVSALNSFGOSPY	600
QY	601	LEFPCTVHLAPKLAVRTPLKAVQAVEGGEVTFSDLTVASAGENFLDGOALKASSVYEH	660
DB	601	LEFPCTVHLAPKLAVRTPLKAVQAVEGGEVTFSDLTVASAGENFLDGOALKASSVYEH	660

QY	661	CDTRHTLTIREVPASLHGAOLKFVANGIESSIRMEVRAAPGLTANKPPAAAAREVLA	720
DB	661	CDTRHTLTIREVPASLHGAOLKFVANGIESSIRMEVRAAPGLTANKPPAAAAREVLA	720
QY	721	HEEAQLLAELSQAQAAVTWLDGRTLSPGPKYEQASAGRRVLLVRDVARDDAGLYECVS	780
DB	721	HEEAQLLAELSQAQAAVTWLDGRTLSPGPKYEQASAGRRVLLVRDVARDDAGLYECVS	780
QY	781	RGGRVIAQLSVQGLARFLHDKMAGSCVDVAVAGGPAQPECETSEAHVHVHVKDGMELGHS	840
DB	781	RGGRVIAQLSVQGLARFLHDKMAGSCVDVAVAGGPAQPECETSEAHVHVHVKDGMELGHS	840
QY	841	GERFLOEDVGRHRLVAATVTRQDEGTYSYCRVGDSDVDVRLRVSEPKVVFVFAKEOLARRKL	900
DB	841	GERFLOEDVGRHRLVAATVTRQDEGTYSYCRVGDSDVDVRLRVSEPKVVFVFAKEOLARRKL	900
QY	901	QAEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVCMHEATGCTRRLLVVQAGQADAGEY	960
DB	901	QAEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVCMHEATGCTRRLLVVQAGQADAGEY	960
QY	961	SCEAGGQRLSPHLVDVKEPKVFAKQVHAHSEVQAEAGANATLSCEVAQAQAEVVMYKDGK	1020
DB	961	SCEAGGQRLSPHLVDVKEPKVFAKQVHAHSEVQAEAGANATLSCEVAQAQAEVVMYKDGK	1020
QY	1021	KLSSSLKVHVEAKGCRRLRVVQAGKTDAGDYSCEARGQVSRFLRHITTEPKMFAKEQSV	1080
DB	1021	KLSSSLKVHVEAKGCRRLRVVQAGKTDAGDYSCEARGQVSRFLRHITTEPKMFAKEQSV	1080
QY	1081	HNEVQAEAGASAMLSCEVAQAQTEVTWYKDGKLSSSSKVGMVEYKGTCTRRLLVLPQAGKAD	1140
DB	1081	HNEVQAEAGASAMLSCEVAQAQTEVTWYKDGKLSSSSKVGMVEYKGTCTRRLLVLPQAGKAD	1140
QY	1141	AGEYSCSAGGQVSRFLHITTEPKGVFAKEQSVHNEVQAEAGTAMLSCEVAQAQTEVTWY	1200
DB	1141	AGEYSCSAGGQVSRFLHITTEPKGVFAKEQSVHNEVQAEAGTAMLSCEVAQAQTEVTWY	1200
QY	1201	KDGKLSSSSKVMEVKGCTRRLLVVQVQKADAGEYSCAGGQVSRFQHLHITTEPKVPAK	1260
DB	1201	KDGKLSSSSKVMEVKGCTRRLLVVQVQKADAGEYSCAGGQVSRFQHLHITTEPKVPAK	1260
QY	1261	EQLVHNEVTRTEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVRIEBAAGCMRQLVQQA	1320
DB	1261	EQLVHNEVTRTEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVRIEBAAGCMRQLVQQA	1320
QY	1321	QOADAGEYTCAGGQRLSPHLVDVSEPKVFAKEQLAHRKVQAEAGATATLSCEVAQAQTE	1380
DB	1321	QOADAGEYTCAGGQRLSPHLVDVSEPKVFAKEQLAHRKVQAEAGATATLSCEVAQAQTE	1380
QY	1381	VTWYKDGKLSSSSKVMEVKGCTRRLLVVQVQKADAGEYSCAGGQRLSPSLDVAEPKV	1440
DB	1381	VTWYKDGKLSSSSKVMEVKGCTRRLLVVQVQKADAGEYSCAGGQRLSPSLDVAEPKV	1440
QY	1441	VFAKEQVPHREVQQAQAGTATLSCEVAQAQTEVTWYKDGKLSSSSKVMEVKGCTRRLLV	1500
DB	1441	VFAKEQVPHREVQQAQAGTATLSCEVAQAQTEVTWYKDGKLSSSSKVMEVKGCTRRLLV	1500
QY	1501	VQQAQADAGEYSCAGGQRLSPHLHVAEPKVPFAKEQASREVQAEAGTATLSCEVAQ	1560
DB	1501	VQQAQADAGEYSCAGGQRLSPHLHVAEPKVPFAKEQASREVQAEAGTATLSCEVAQ	1560
QY	1561	AQTEVTWYKDGKLSSSSKVMEVKGCTRRLLVVQAEQADAGEYSCKAGQRLSPFLHVA	1620
DB	1561	AQTEVTWYKDGKLSSSSKVMEVKGCTRRLLVVQAEQADAGEYSCKAGQRLSPFLHVA	1620
QY	1621	EPKVVFAKEQVPHREVQAEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVMEVKGCT	1680
DB	1621	EPKVVFAKEQVPHREVQAEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVMEVKGCT	1680
QY	1681	RRLVVQAGQADAGEYSCAGGQRLSPFLHVAEIPQISERPCCRREPLVVVKEHEDITLTA	1740
DB	1681	RRLVVQAGQADAGEYSCAGGQRLSPFLHVAEIPQISERPCCRREPLVVVKEHEDITLTA	1740

QY 1741 TLATPSAATVTLKDGVEIRRSKGHETASOGDHTLTVHGAQVLDIAIYSCRVBAGQDP 1800
DB 1741 TLATPSAATVTLKDGVEIRRSKGHETASOGDHTLTVHGAQVLDIAIYSCRVBAGQDP 1800
QY 1801 PVQVEEVAAKFRLLEPVCGLGGTTLTACELSPACAEVVRWCGNTQPRVGRKFQWVABG 1860
DB 1801 PVQVEEVAAKFRLLEPVCGLGGTTLTACELSPACAEVVRWCGNTQPRVGRKFQWVABG 1860
QY 1861 PVRSLTVLGLRAEDAGEYCESRDDHTSAQLTVSVPRVVKFMSGLSTVVAEBEGEATPQC 1920
DB 1861 PVRSLTVLGLRAEDAGEYCESRDDHTSAQLTVSVPRVVKFMSGLSTVVAEBEGEATPQC 1920
QY 1921 VVPSDVAVVRPDDGALLQPSKEPAISQSGASHSLTISDLVLEDAGQIITVEABEGSSAA 1980
DB 1921 VVPSDVAVVRPDDGALLQPSKEPAISQSGASHSLTISDLVLEDAGQIITVEABEGSSAA 1980
QY 1981 LRVREAPVLFKKKJLPQTVEERSSVTVLELTPMPPELRTWNTATNALAPKKNVEHAEGA 2040
DB 1981 LRVREAPVLFKKKJLPQTVEERSSVTVLELTPMPPELRTWNTATNALAPKKNVEHAEGA 2040
QY 2041 RHRLVLHNVGFPADRGFFGCETPDDKTQAKLTVMERQVRLVRLGLQAVEAREOQTATMEVOL 2100
DB 2041 RHRLVLHNVGFPADRGFFGCETPDDKTQAKLTVMERQVRLVRLGLQAVEAREOQTATMEVOL 2100
QY 2101 SHADVGSWTRDGLRFQOQPTCHLAVRGPMMHTLTLISGLRPEDSGLMVFKAEGVHTSARLV 2160
DB 2101 SHADVGSWTRDGLRFQOQPTCHLAVRGPMMHTLTLISGLRPEDSGLMVFKAEGVHTSARLV 2160
QY 2161 VTLEPVSFSRPLQDVVVTTKEKVTLCELSRPNVDVRLKDGVELRAGKTMALAAQACR 2220
DB 2161 VTLEPVSFSRPLQDVVVTTKEKVTLCELSRPNVDVRLKDGVELRAGKTMALAAQACR 2220
QY 2221 SLTIYRCBPADGVVCDADHAQSSASVKVQGRTYTLIYRRVLAEDAGEIQFVAENABSR 2280
DB 2221 SLTIYRCBPADGVVCDADHAQSSASVKVQGRTYTLIYRRVLAEDAGEIQFVAENABSR 2280
QY 2281 AOLRVKELPVTILVRLDKIANEKHGVLECOVSRASAQVRWFKGSOELQPKYELVSD 2340
DB 2281 AOLRVKELPVTILVRLDKIANEKHGVLECOVSRASAQVRWFKGSOELQPKYELVSD 2340
QY 2341 GLYRLKLIISDVHAEDEDTYTCADGVKTSQAQFFVEEQSIITVIRGLQDVTVMPEPAPWPEC 2400
DB 2341 GLYRLKLIISDVHAEDEDTYTCADGVKTSQAQFFVEEQSIITVIRGLQDVTVMPEPAPWPEC 2400
QY 2401 ETSIPSVRPPKWLKGLTKVLQAGNVGLEQEGTVHRLMLRRTCTWTGPHVFTVGRSSA 2460
DB 2401 ETSIPSVRPPKWLKGLTKVLQAGNVGLEQEGTVHRLMLRRTCTWTGPHVFTVGRSSA 2460
QY 2461 RLWVSDIPVLTFRPLEPKTGRELOQSVLSCDPRPAPKAVQWYKDDTPLSPSEKFKMSLEG 2520
DB 2461 RLWVSDIPVLTFRPLEPKTGRELOQSVLSCDPRPAPKAVQWYKDDTPLSPSEKFKMSLEG 2520
QY 2521 QMAELRILRLMPADAGVRCQAGSAHSTEVTVAREVTVTGLODAEATEEGMASFSCE 2580
DB 2521 QMAELRILRLMPADAGVRCQAGSAHSTEVTVAREVTVTGLODAEATEEGMASFSCE 2580
QY 2581 LSHDEEVEWSLNGMPLYNDSFHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTARS 2640
DB 2581 LSHDEEVEWSLNGMPLYNDSFHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTARS 2640
QY 2641 EVRVKPVVFLKALDLSAEERTLALQCEVSDPEAHVVRWKGVLQSPDKYDFLHTAGT 2700
DB 2641 EVRVKPVVFLKALDLSAEERTLALQCEVSDPEAHVVRWKGVLQSPDKYDFLHTAGT 2700
QY 2701 RGLVHVDVSPEDAGLYTCHVGSEETRARVRVHDLHVGITKRLKTMVELEGESCSPECVLS 2760
DB 2701 RGLVHVDVSPEDAGLYTCHVGSEETRARVRVHDLHVGITKRLKTMVELEGESCSPECVLS 2760
QY 2761 HESADPAMTVGGKTVGSSSRFQATROGRKYILVVRREAPSDAGEVFSVRGLTSKASL 2820
DB 2761 HESADPAMTVGGKTVGSSSRFQATROGRKYILVVRREAPSDAGEVFSVRGLTSKASL 2820
QY 2821 IVRERPAAIIPKLEDPQWAPGEDVELRCELRSAGTFPVHMLKDKRAIKRSQKYDVVCEGTM 2880

DB 2821 IVRERPAAIIPKLEDPQWAPGEDVELRCELRSAGTFPVHMLKDKRAIKRSQKYDVVCEGTM 2880
QY 2881 AMLVIRGASLKDAGEYTCVEASKSTASLHVBEKANCFTTEELTNLQVEBKGTAVFTCKTE 2940
DB 2881 AMLVIRGASLKDAGEYTCVEASKSTASLHVBEKANCFTTEELTNLQVEBKGTAVFTCKTE 2940
QY 2941 HPAATVTVWRKGLLELRASGHQPSQEGTLRLTITISALEKADSDTYTCDIQOASRAQLLV 3000
DB 2941 HPAATVTVWRKGLLELRASGHQPSQEGTLRLTITISALEKADSDTYTCDIQOASRAQLLV 3000
QY 3001 QGRRVHIIIEDLVDOVBQESSATFCRISPNANYPVHMFDTPLHANELNEIDAOPGGY 3060
DB 3001 QGRRVHIIIEDLVDOVBQESSATFCRISPNANYPVHMFDTPLHANELNEIDAOPGGY 3060
QY 3061 HVLTIRLQALKDSGTIYFEAGDQASAAALRVTEKPSVFSRELTDATITEGEDLTLCVETS 3120
DB 3061 HVLTIRLQALKDSGTIYFEAGDQASAAALRVTEKPSVFSRELTDATITEGEDLTLCVETS 3120
QY 3121 TCDIPMCWTGDKTLURGSARCOLSHEGHRAQLLITGATLQDSGRYKCEAGGACSSIVRV 3180
DB 3121 TCDIPMCWTGDKTLURGSARCOLSHEGHRAQLLITGATLQDSGRYKCEAGGACSSIVRV 3180
QY 3181 HARPVRFQEALKDLEVLGGATLRCVLSVAAPVKWCYGNVLRPGDKYSLROEGAMLE 3240
DB 3181 HARPVRFQEALKDLEVLGGATLRCVLSVAAPVKWCYGNVLRPGDKYSLROEGAMLE 3240
QY 3241 LVVRLRLPQDSGRYSCSGDQTTSATLTVTALPAQFIKGLRKEATEGATATLRCBSLKT 3300
DB 3241 LVVRLRLPQDSGRYSCSGDQTTSATLTVTALPAQFIKGLRKEATEGATATLRCBSLKT 3300
QY 3301 APVWRKGSSETLRDGDYCLRDQDGMCELOIRGLAMVDAAEYSCVCGBERTSASLTIRPM 3360
DB 3301 APVWRKGSSETLRDGDYCLRDQDGMCELOIRGLAMVDAAEYSCVCGBERTSASLTIRPM 3360
QY 3361 PAHFTIGRLRHQESIEGATATLRCBSLKAAPVWRKGRSLRDGDHSLRDQDGAVALCLOTC 3420
DB 3361 PAHFTIGRLRHQESIEGATATLRCBSLKAAPVWRKGRSLRDGDHSLRDQDGAVALCLOTC 3420
QY 3421 GLAVADAGEYSCVCGBERTSATLTVKALPAKFTGELRNEEAVEGATAMLCBSLKVAPVE 3480
DB 3421 GLAVADAGEYSCVCGBERTSATLTVKALPAKFTGELRNEEAVEGATAMLCBSLKVAPVE 3480
QY 3481 WRKGPENLRDGDYILRQEGTRCELOICGLAMADAGEYLCVCGBERTSATLTIRALPARF 3540
DB 3481 WRKGPENLRDGDYILRQEGTRCELOICGLAMADAGEYLCVCGBERTSATLTIRALPARF 3540
QY 3541 IEDVKNOQAREGATVLOCELNSAAPVWRKGSSETLRDGDYSLRDQGTCKELOIRGLAM 3600
DB 3541 IEDVKNOQAREGATVLOCELNSAAPVWRKGSSETLRDGDYSLRDQGTCKELOIRGLAM 3600
QY 3601 ADTGEYSCVCGBERTSAMLTVRALPIKFTGELRNEEAEGATAVLRCBSLKNAPVWVKWG 3660
DB 3601 ADTGEYSCVCGBERTSAMLTVRALPIKFTGELRNEEAEGATAVLRCBSLKNAPVWVKWG 3660
QY 3661 HETLRDGDHSLRDQDGMCELOIRGLVAEDAGEYLCMCKERTSAMLTVRAMPSPFIEGL 3720
DB 3661 HETLRDGDHSLRDQDGMCELOIRGLVAEDAGEYLCMCKERTSAMLTVRAMPSPFIEGL 3720
QY 3721 RNEEATEGDTATLWCELSKAAPVWRKGHETLRDGDHSLRDQDGRCELOIRGLAVVDAG 3780
DB 3721 RNEEATEGDTATLWCELSKAAPVWRKGHETLRDGDHSLRDQDGRCELOIRGLAVVDAG 3780
QY 3781 EYSCVCGBERTSATLTVRALPARFIEDVKNOQAREGATVLOCELSKAAPVWVRKGSSETL 3840
DB 3781 EYSCVCGBERTSATLTVRALPARFIEDVKNOQAREGATVLOCELSKAAPVWVRKGSSETL 3840
QY 3841 RGGDRYSLRDQGTTRCELOIHGLSVADTGEYSCVCGBERTSATLTVRAPQVPREPQLSQ 3900
DB 3841 RGGDRYSLRDQGTTRCELOIHGLSVADTGEYSCVCGBERTSATLTVRAPQVPREPQLSQ 3900
QY 3901 ABEGSTATLQCELSSEPTATVWVSGGLQLOQNGRRPRELQGTAEVLQDLOREDTGEYT 3960

Db 3901 ABEGSTATLOCELSEPTATVWWSKGLQLOANGRRPRLOQCTAELVLQDLQREDTGEYT 3960
Qy 3961 CTCISOATSLTTTAAAPVRELRELQHOVEDGTAHLCCELSRAGASVWRKGSLOLFP 4020
Db 3961 CTCISOATSLTTTAAAPVRELRELQHOVEDGTAHLCCELSRAGASVWRKGSLOLFP 4020
Qy 4021 CAKQOMVQDGAABELLVRGVEQEDAGDYTCDTGHTQSMASLSVRPRPKFTRILQSLQE 4080
Db 4021 CAKQOMVQDGAABELLVRGVEQEDAGDYTCDTGHTQSMASLSVRPRPKFTRILQSLQE 4080
Qy 4081 TGDIALRCOLSDAESGAVQWMLKEGVELHAGPKYEMRSQATRELLIHOLEAKDTGEYA 4140
Db 4081 TGDIALRCOLSDAESGAVQWMLKEGVELHAGPKYEMRSQATRELLIHOLEAKDTGEYA 4140
Qy 4141 CVTGGOKTAAASLRVTEPEVTIVRGLVDAEVTADREDFESCVRAGATGVOMCLOGLPLQ 4200
Db 4141 CVTGGOKTAAASLRVTEPEVTIVRGLVDAEVTADREDFESCVRAGATGVOMCLOGLPLQ 4200
Qy 4201 SNEVTEVAVRDRGRIHTLRLKGVTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPLODVQ 4260
Db 4201 SNEVTEVAVRDRGRIHTLRLKGVTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPLODVQ 4260
Qy 4261 LSEGODASFQRLSRASQOEARMALGGVPLQANEMNDITVQGTLLHLLTHKVTLEDAGT 4320
Db 4261 LSEGODASFQRLSRASQOEARMALGGVPLQANEMNDITVQGTLLHLLTHKVTLEDAGT 4320
Qy 4321 VSFHVGTCCSSAOLKVTAKNTVVRGLENVAELEGEALFECQLOPEVAAHTWLLDDEPV 4380
Db 4321 VSFHVGTCCSSAOLKVTAKNTVVRGLENVAELEGEALFECQLOPEVAAHTWLLDDEPV 4380
Qy 4381 RTSNAEVVFPENGRLHLLKLNLRPQDSCRVTFLAGDMVTSAFILTVRGMREILEPLKN 4440
Db 4381 RTSNAEVVFPENGRLHLLKLNLRPQDSCRVTFLAGDMVTSAFILTVRGMREILEPLKN 4440
Qy 4441 AAVRAGAQAQRTCTLSEAVPVGEASWYINGAAVOPDDSDWTADTADGSHQALLLSAQPHH 4500
Db 4441 AAVRAGAQAQRTCTLSEAVPVGEASWYINGAAVOPDDSDWTADTADGSHQALLLSAQPHH 4500
Qy 4501 AGEVTFACRDVAVASARLTVLGLPDDPEAEVVVAHSSHVTLSWAAPMSDGGGLCGYRVE 4560
Db 4501 AGEVTFACRDVAVASARLTVLGLPDDPEAEVVVAHSSHVTLSWAAPMSDGGGLCGYRVE 4560
Qy 4561 VKEGATQWRLCHELVPGPECVWDGLAPGETYRFRVAAVGPVGAGEPVHLPTQTVRLAEP 4620
Db 4561 VKEGATQWRLCHELVPGPECVWDGLAPGETYRFRVAAVGPVGAGEPVHLPTQTVRLAEP 4620
Qy 4621 KPVPPOPSAPESRQVAAGEDVLSLEVVAEAGEVIVHKGMERIQPGGRFVVSQGRQML 4680
Db 4621 KPVPPOPSAPESRQVAAGEDVLSLEVVAEAGEVIVHKGMERIQPGGRFVVSQGRQML 4680
Qy 4681 VIKGFTAEDQGEYHCGLAQSGISCPAATFQVALSPASVDEAPQPSLPPEAAQEGDLHLW 4740
Db 4681 VIKGFTAEDQGEYHCGLAQSGISCPAATFQVALSPASVDEAPQPSLPPEAAQEGDLHLW 4740
Qy 4741 EALARKRMSREPTLDSISELPEEDGRSQRLPQEAEEVAPDLSEGYSTADSLARTGDADL 4800
Db 4741 EALARKRMSREPTLDSISELPEEDGRSQRLPQEAEEVAPDLSEGYSTADSLARTGDADL 4800
Qy 4801 SHTSDSDSRAGTPSLVTLKAGBPGTSPASKVGAPAAAPSVKPOQOEPPLAARVPLG 4860
Db 4801 SHTSDSDSRAGTPSLVTLKAGBPGTSPASKVGAPAAAPSVKPOQOEPPLAARVPLG 4860
Qy 4861 DLSTKDLGDPSMDKAANKVIAAFKGYKVRKEMKQOEGPMFHTFGDTAQVGDALRLFCV 4920
Db 4861 DLSTKDLGDPSMDKAANKVIAAFKGYKVRKEMKQOEGPMFHTFGDTAQVGDALRLFCV 4920
Qy 4921 VASKADVPAARWLKDGVELTDRGHHHIDQLDGTCSLLIAGLDRADAGCYTCQVSNKFGQV 4980
Db 4921 VASKADVPAARWLKDGVELTDRGHHHIDQLDGTCSLLIAGLDRADAGCYTCQVSNKFGQV 4980
Qy 4981 THSACVVVSGSESEASESSGGELODAFRRAARLHRLFRTPSPAEVSDEELFLSADEGPA 5040
Db 4981 THSACVVVSGSESEASESSGGELODAFRRAARLHRLFRTPSPAEVSDEELFLSADEGPA 5040

Qy 5041 EPBEPADQWYRDEHPICIRFEALTEARQAVTRFQEMFATLIGIVBEIKLVEOQPRVEM 5100
Db 5041 EPBEPADQWYRDEHPICIRFEALTEARQAVTRFQEMFATLIGIVBEIKLVEOQPRVEM 5100
Qy 5101 CISKETPAPVPPPEPLPSLLTSDAAPVFLTELQNEQVODGYPVSDCVVVGQPMPSVRWF 5160
Db 5101 CISKETPAPVPPPEPLPSLLTSDAAPVFLTELQNEQVODGYPVSDCVVVGQPMPSVRWF 5160
Qy 5161 KDGKLLBEDDHYMINEDQOGGHOLIITAVVPADMVGVYRCLAENSMGVSSYTKAEULRVDLTS 5220
Db 5161 KDGKLLBEDDHYMINEDQOGGHOLIITAVVPADMVGVYRCLAENSMGVSSYTKAEULRVDLTS 5220
Qy 5221 TDYDTAADATSESSYFSAQGYLSRREQGETSTTDEGQLPQWVEELRDLQVAPQTRIAKF 5280
Db 5221 TDYDTAADATSESSYFSAQGYLSRREQGETSTTDEGQLPQWVEELRDLQVAPQTRIAKF 5280
Qy 5281 QLKVKGYPAARLWFKDGOPLTASAHIRMTGKKILHLEIISVTRRESGGYAAVISNAMG 5340
Db 5281 QLKVKGYPAARLWFKDGOPLTASAHIRMTGKKILHLEIISVTRRESGGYAAVISNAMG 5340
Qy 5341 AAYSSARLLVRGPDPEPEKPADVHEQLVPPMLERFTPKVKKGSSITTSVKVEGRPV 5400
Db 5341 AAYSSARLLVRGPDPEPEKPADVHEQLVPPMLERFTPKVKKGSSITTSVKVEGRPV 5400
Qy 5401 TVHWLREAEARGVLWIGPDTPGYTVVASSAQOHSI.VLLDVGRQHOGTYTCTIASNAAGQALC 5460
Db 5401 TVHWLREAEARGVLWIGPDTPGYTVVASSAQOHSI.VLLDVGRQHOGTYTCTIASNAAGQALC 5460
Qy 5461 SASLHVSGLPKVBEOEKVKEALISTFLOGTQOASIAOGLTASPADIGGQKEPLAAKE 5520
Db 5461 SASLHVSGLPKVBEOEKVKEALISTFLOGTQOASIAOGLTASPADIGGQKEPLAAKE 5520
Qy 5521 ALGHLSLAEVTEBEFLQKLSQITEMVSAKITQAKLOVPGGDSDEDSKTSPASPRHGRSR 5580
Db 5521 ALGHLSLAEVTEBEFLQKLSQITEMVSAKITQAKLOVPGGDSDEDSKTSPASPRHGRSR 5580
Qy 5581 PSSSIOESSSESDGDARGEIFDIYVVTADYLP.LGABQDAITTLREGQYVVEVLDAAHPLRW 5640
Db 5581 PSSSIOESSSESDGDARGEIFDIYVVTADYLP.LGABQDAITTLREGQYVVEVLDAAHPLRW 5640
Qy 5641 LVRTKPTKSPSGQWSPAYLDRRLKLSPEWGAABAEPEPGEAVSEDEYKARLSSVIOE 5700
Db 5641 LVRTKPTKSPSGQWSPAYLDRRLKLSPEWGAABAEPEPGEAVSEDEYKARLSSVIOE 5700
Qy 5701 LLSSEQAFVBELOFLOSHHLQHLERCPHPTAVAGQKAVIFRNVRDIGRPHSSFLQELQQ 5760
Db 5701 LLSSEQAFVBELOFLOSHHLQHLERCPHPTAVAGQKAVIFRNVRDIGRPHSSFLQELQQ 5760
Qy 5761 CDTDDDDVAMCFIKNQAAFEQYLEFLVGRVQAESVVVSTAIQEFYKKYABEALLAGDPSQP 5820
Db 5761 CDTDDDDVAMCFIKNQAAFEQYLEFLVGRVQAESVVVSTAIQEFYKKYABEALLAGDPSQP 5820
Qy 5821 PPPLOHYLPOPVERVQRYQALLKELIRNKARNQNCALLBQAAVAVSALPQRAENKLHV 5880
Db 5821 PPPLOHYLPOPVERVQRYQALLKELIRNKARNQNCALLBQAAVAVSALPQRAENKLHV 5880
Qy 5881 SLMENYPTGLEALCEPIRQGHFIVWEGAPGARMFWKGNHRVFLFRNHLVICKPRRDSRT 5940
Db 5881 SLMENYPTGLEALCEPIRQGHFIVWEGAPGARMFWKGNHRVFLFRNHLVICKPRRDSRT 5940
Qy 5941 DTVSIVFRNMWMLSSIDLNDQVEGDDRAFEVWQGRESVRKYLLOARTAIKSSWVKIC 6000
Db 5941 DTVSIVFRNMWMLSSIDLNDQVEGDDRAFEVWQGRESVRKYLLOARTAIKSSWVKIC 6000
Qy 6001 GIQORLALPVWRPDPFEEELADCTABELGETVKLACRVGTGTPKPVISWYKDGKAVQVDPHH 6060
Db 6001 GIQORLALPVWRPDPFEEELADCTABELGETVKLACRVGTGTPKPVISWYKDGKAVQVDPHH 6060
Qy 6061 ILIEDPDGSCALILDSLTGVDSDGQYMCFAASAGNCSTLGKILVQVPPRFVNVKVRASPFV 6120
Db 6061 ILIEDPDGSCALILDSLTGVDSDGQYMCFAASAGNCSTLGKILVQVPPRFVNVKVRASPFV 6120

QY 6121 EGEDAQFTCTIEGAPYPOIRWYKOGALLTGNKFTQLSEPRSGLLVLVIRAAKSDGLGY 6180
DB 6121 EGEDAQFTCTIEGAPYPOIRWYKOGALLTGNKFTQLSEPRSGLLVLVIRAAKSDGLGY 6180
QY 6181 ECELVNRLGSRASASAELEI QSPMLCAQOCHREQLVAAVEDTTLERADQEVTSVLKRLLG 6240
DB 6181 ECELVNRLGSRASASAELEI QSPMLCAQOCHREQLVAAVEDTTLERADQEVTSVLKRLLG 6240
QY 6241 PKAPGPGTGLTGPFCPRGAPALQETGSPVPTCTSEAPVPPVPPVPOPLLHHEGPEQPE 6300
DB 6241 PKAPGPGTGLTGPFCPRGAPALQETGSPVPTCTSEAPVPPVPPVPOPLLHHEGPEQPE 6300
QY 6301 AIARAQEWTVFIRMEGAAPGAGTCELLWDVHSHVVRRETTQRTYTYQAIDHTHARPPSMQ 6360
DB 6301 AIARAQEWTVFIRMEGAAPGAGTCELLWDVHSHVVRRETTQRTYTYQAIDHTHARPPSMQ 6360
QY 6361 VTIEDVQAOTGTAQFEAI IEGDPOPSVTWYKDSVQLVDSFTRLSQQOEGTTVSLVLRHVA 6420
DB 6361 VTIEDVQAOTGTAQFEAI IEGDPOPSVTWYKDSVQLVDSFTRLSQQOEGTTVSLVLRHVA 6420
QY 6421 SKDAGVYTCLAQNTGGQVLCBAELLVLGDNEDPDESEKSHRRKLSHFYEVKEEIGRGVFG 6480
DB 6421 SKDAGVYTCLAQNTGGQVLCBAELLVLGDNEDPDESEKSHRRKLSHFYEVKEEIGRGVFG 6480
QY 6481 FVKRVQHGKNIKLAACFIPLRSRTAQAYRERDILAALSHPLVTGLLDQFETRKTILII 6540
DB 6481 FVKRVQHGKNIKLAACFIPLRSRTAQAYRERDILAALSHPLVTGLLDQFETRKTILII 6540
QY 6541 LELCSSEBLLDRLYRKGVVTAEVKVIQQLVEGLHYLHSHGVLDI KPSNLMVHPAR 6600
DB 6541 LELCSSEBLLDRLYRKGVVTAEVKVIQQLVEGLHYLHSHGVLDI KPSNLMVHPAR 6600
QY 6601 EDIKICDFGAQNTIPPAELQFSQYSGSPFVSPETIQONPVSEASDIWANGVISYLSLTC 6660
DB 6601 EDIKICDFGAQNTIPPAELQFSQYSGSPFVSPETIQONPVSEASDIWANGVISYLSLTC 6660
QY 6661 SPPAGESDRATLLNVLEGRVSWSPMAAHLSEDAKDFIKATLQAPAPRPSAAQCLSHPW 6720
DB 6661 SPPAGESDRATLLNVLEGRVSWSPMAAHLSEDAKDFIKATLQAPAPRPSAAQCLSHPW 6720
QY 6721 FLKSNPABEAHFIPTQKFLARSRWORSLSYKSIILVMRSIPELLRGPDPSPSLGVAR 6780
DB 6721 FLKSNPABEAHFIPTQKFLARSRWORSLSYKSIILVMRSIPELLRGPDPSPSLGVAR 6780
QY 6781 HLCRDGTGSSSSSSSDNELAPFARAKSLPSPVTHSPLLPGRFLRPSASLPEAEASE 6840
DB 6781 HLCRDGTGSSSSSSSDNELAPFARAKSLPSPVTHSPLLPGRFLRPSASLPEAEASE 6840
QY 6841 RSTAPAPAPSPGAGPAAQGCVPRHVSIRSLFYHQAGESPEHGALAPGSRHRHPARRH 6900
DB 6841 RSTAPAPAPSPGAGPAAQGCVPRHVSIRSLFYHQAGESPEHGALAPGSRHRHPARRH 6900
QY 6901 LLKGYIAGALPGLREPLMEHRVLEEEAREEQATLLAKAPSFETALRIPASGTHLAPGH 6960
DB 6901 LLKGYIAGALPGLREPLMEHRVLEEEAREEQATLLAKAPSFETALRIPASGTHLAPGH 6960
QY 6961 SHSLEHDSPTSPRPSSEACGAQLRPSAPSGAPIRDMGHQSGKQLPSTGCHPGTAQPE 7020
DB 6961 SHSLEHDSPTSPRPSSEACGAQLRPSAPSGAPIRDMGHQSGKQLPSTGCHPGTAQPE 7020
QY 7021 RPSDPSPMGQAPFCHPKQGSAPQGGCSPHPAVAPCPGSPFPFGCKEAPLVPSSPFLGQ 7080
DB 7021 RPSDPSPMGQAPFCHPKQGSAPQGGCSPHPAVAPCPGSPFPFGCKEAPLVPSSPFLGQ 7080
QY 7081 POAPAPAKASPPPLDSKMGPGDISLPGRPKGPCSSPGCSASQSSQVSSLRVGSQVGT 7140
DB 7081 POAPAPAKASPPPLDSKMGPGDISLPGRPKGPCSSPGCSASQSSQVSSLRVGSQVGT 7140
QY 7141 EPGPSLDAEGWTQAEADLSDTPTLQRPQEOATMRKFSLGGRGVAGVAGYGTTFAGGDA 7200
DB 7141 EPGPSLDAEGWTQAEADLSDTPTLQRPQEOATMRKFSLGGRGVAGVAGYGTTFAGGDA 7200
QY 7201 GGMGLQGPMWARIAWAVSQSEEEQEARAESQSEEQEARAESPLQVARSARVPVEVGRA 7260

DB 7201 GGMGLQGPMWARIAWAVSQSEEEQEARAESQSEEQEARAESPLQVARSARVPVEVGRA 7260
QY 7261 PTRSSPEPTPWEDIGQVSLVQIRLSDGDAEADTISLDISEVDPAYLNLSDLYDKYLPF 7320
DB 7261 PTRSSPEPTPWEDIGQVSLVQIRLSDGDAEADTISLDISEVDPAYLNLSDLYDKYLPF 7320
QY 7321 EFMIFRKVPKSAQPEPPSPMAEELAEPEPTWMPGELGPHAGLEITEESSEDVALLAE 7380
DB 7321 EFMIFRKVPKSAQPEPPSPMAEELAEPEPTWMPGELGPHAGLEITEESSEDVALLAE 7380
QY 7381 AAVGRKRWSSPSRSLPHFGRHLPLDPAELGLRERVKASVEHISRLIKGRPEGLEKEG 7440
DB 7381 AAVGRKRWSSPSRSLPHFGRHLPLDPAELGLRERVKASVEHISRLIKGRPEGLEKEG 7440
QY 7441 PPRKKPGLASFRLSGLKSWDRAPTFLRELSDETIVLGOSVTLACQVSAQPAQAATWSKDG 7500
DB 7441 PPRKKPGLASFRLSGLKSWDRAPTFLRELSDETIVLGOSVTLACQVSAQPAQAATWSKDG 7500
QY 7501 APLSSSSRVLISATLKNFQLLTILVVAEDLGVTCTSVSNALGTVTITGVLRKAERPSS 7560
DB 7501 APLSSSSRVLISATLKNFQLLTILVVAEDLGVTCTSVSNALGTVTITGVLRKAERPSS 7560
QY 7561 PCPDIGEVYADGVLLWVKPVSYPVTYIVOCSEGGSWTTLASDIFDCCYLTSLSRGG 7620
DB 7561 PCPDIGEVYADGVLLWVKPVSYPVTYIVOCSEGGSWTTLASDIFDCCYLTSLSRGG 7620
QY 7621 TYTFRACVSKAGMGPYSSPSEQVLLGGPSHLASEEESQGRSAQPLPSTKTFAPQTQIOR 7680
DB 7621 TYTFRACVSKAGMGPYSSPSEQVLLGGPSHLASEEESQGRSAQPLPSTKTFAPQTQIOR 7680
QY 7681 GRFSVVRQCEKASGRALAAKIIIPYHPKDKTAVLREYBALKGLRHPHLAQLHAAVLSPRH 7740
DB 7681 GRFSVVRQCEKASGRALAAKIIIPYHPKDKTAVLREYBALKGLRHPHLAQLHAAVLSPRH 7740
QY 7741 LVLILELCSGPELLPCLAEARASYSESEVKDYLWMLSATQYLHNOHILHDLRSENMIIT 7800
DB 7741 LVLILELCSGPELLPCLAEARASYSESEVKDYLWMLSATQYLHNOHILHDLRSENMIIT 7800
QY 7801 EYNLLKVVDLGNASLSQEKVLPDKFYDLETWAPLELGGQVAPQTDIWAIGVTAFTM 7860
DB 7801 EYNLLKVVDLGNASLSQEKVLPDKFYDLETWAPLELGGQVAPQTDIWAIGVTAFTM 7860
QY 7861 LSAYPVSSSEGARDLQRLKGLVRLSRCYAGLSGAVAFRLSTLCAQPWGRPCASSCLQ 7920
DB 7861 LSAYPVSSSEGARDLQRLKGLVRLSRCYAGLSGAVAFRLSTLCAQPWGRPCASSCLQ 7920
QY 7921 CPWLTEEGPACSRPAPVTFTPTARLRFVVRNREKRALLYKRNHNAQVR 7968
DB 7921 CPWLTEEGPACSRPAPVTFTPTARLRFVVRNREKRALLYKRNHNAQVR 7968
RESULT 2
ADJ70485
ID ADJ70485 standard; protein; 6620 AA.
XX
AC ADJ70485;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human heat mitochondrial protein as a therapeutic target SeqID2291.
XX
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX
OS Homo sapiens.
XX
PN WO2003087768-A2.

XX 23-OCT-2003.
XX 04-APR-2003; 2003WO-US010870.
XX 12-APR-2002; 2002US-0372843P.
XX 17-JUN-2002; 2002US-0389987P.
XX 20-SEP-2002; 2002US-0412418P.
XX (MITO-) MITOKOR.
XX (BUCK-) BUCK INST AGE RES.
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
XX Warnock DB;
XX WPI; 2003-845369/78.
XX Identifying a mitochondrial target for drug screening assays and for
XX treating diseases associated with altered mitochondrial function,
XX comprises detecting a modified polypeptide in a sample and correlating
XX with the disease.
XX Claim 1; SEQ ID NO 2291; 180pp; English.
XX This invention relates to novel mitochondrial targets that can be used
XX for therapeutic intervention in treating a disease associated with
XX altered mitochondrial function. Specifically, it refers to a method for
XX identifying proteins of the human heart mitochondrial proteome that are
XX useful for drug screening assays, as well as therapeutic targets. The
XX present invention describes a method for identifying such proteins that
XX can be used in the treatment of various diseases associated with altered
XX mitochondrial function including diabetes mellitus, Huntington's disease,
XX osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
XX cephalopathy, lactic acidosis and stroke (MELAS), myoclonic epilepsy
XX ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
XX compositions have neuroprotective, neurotropic, antidiabetic,
XX anticonvulsant, antiarthritic, osteopathic, ophthalmological and
XX cytoskeletal activities. This polypeptide sequence is a human heart
XX mitochondrial protein of the invention.
XX Sequence 6620 AA;
Query Match 77.9%; Score 32134; DB 7; Length 6620;
Best Local Similarity 98.1%; Pred. NO. 0;
Matches 6250; Conservative 15; Mismatches 57; Indels 48; Gaps 9;
Qy 1 MDQPFSGAPRFLTRPKAFVSVGKDATLSQIVGNPTPOVSWEKDQOPVTAGARFLAQ 60
Db 1 MDQPFSGAPRFLTRPKAFVSVGKDATLSQIVGNPTPOVSWEKDQOPVTAGARFLAQ 60
Qy 61 DGDLYRLTILDALGDSQYVCRARNALGEAFVAVGLQVDAEAAEQAHPFLRLPTSIR 120
Db 61 DGDLYRLTILDALGDSQYVCRARNALGEAFVAVGLQVDAEAAEQAHPFLRLPTSIR 120
Qy 121 VREGSEATFRCKVGSPPRPAVSWSKDGRRLGEPDGPVRVVEELGEASALRIRAAAPRDGG 180
Db 121 VREGSEATFRCKVGSPPRPAVSWSKDGRRLGEPDGPVRVVEELGEASALRIRAAAPRDGG 180
Qy 181 TYEVAENPLGAASAAALVDSDAADTASPRGSTAALLAHLQRRREMAEGAPASPP 240
Db 181 TYEVAENPLGAASAAALVDSDAADTASPRGSTAALLAHLQRRREMAEGAPASPP 240
Qy 241 STGTRTCTVTGKHLARLSCTYTGPKPTVMKKDQQLVTEGRRHVVDQAENFVLKILF 300
Db 241 STGTRTCTVTGKHLARLSCTYTGPKPTVMKKDQQLVTEGRRHVVDQAENFVLKILF 300
Qy 301 CKQSDRGLYTCTASNLVGQYTSVSVLVVVRBPAVFPKRLQDLVREKESATFLCEVPQPS 360
Db 301 CKQSDRGLYTCTASNLVGQYTSVSVLVVVRBPAVFPKRLQDLVREKESATFLCEVPQPS 360
Qy 361 TEAAWFKETRLWASAKYGIIEEGTERLTVRNVSDDDAVYICETPEGSRVLAELAVQG 420
Db 361 TEAAWFKETRLWASAKYGIIEEGTERLTVRNVSDDDAVYICETPEGSRVLAELAVQG 420

Qy 421 NLLRKLPRKTAVRVGDGTAMFCVELAVPVPVHVLNRNQEEVVAGGRVAISAEGRHTLTIS 480
Db 421 NLLRKLPRKTAVRVGDGTAMFCVELAVPVPVHVLNRNQEEVVAGGRVAISAEGRHTLTIS 480
Qy 481 QCCLLEDVGVQAFMAGDCQSTSTRFCVSAPRRPPLQPPVDPVVKARMESVILSNPPPHGE 540
Db 481 QCCLLEDVGVQAFMAGDCQSTSTRFCVSAPRRPPLQPPVDPVVKARMESVILSNPPPHGE 540
Qy 541 RPTVIDGVLVEKKLGTYYTWIRCHEAEWATPELTVDVABEGNFQFRVSNALNFGQSPY 600
Db 541 RPTVIDGVLVEKKLGTYYTWIRCHEAEWATPELTVDVABEGNFQFRVSNALNFGQSPY 600
Qy 601 LEFPGTVHLAPKLAVRTPLKAVQAVEGGEVTFSDLTVASAGEWFLDQALKASSVYBIH 660
Db 601 LEFPGTVHLAPKLAVRTPLKAVQAVEGGEVTFSDLTVASAGEWFLDQALKASSVYBIH 660
Qy 661 CDRTRHTLTITREVPASLHGALQKPVANGIESSIRMEVRAAPCLGTANKPAAAAAEVLARL 720
Db 661 CDRTRHTLTITREVPASLHGALQKPVANGIESSIRMEVRAAPCLGTANKPAAAAAEVLARL 720
Qy 721 HEEAQLLAELSDQAAAVTWLKGRTLSGPKPYEQVQASAGRRVLLVRDVARDDAGLYECVS 780
Db 721 HEEAQLLAELSDQAAAVTWLKGRTLSGPKPYEQVQASAGRRVLLVRDVARDDAGLYECVS 780
Qy 781 RGGRIAYQLSVQGLARFLHKDMAGSCYDVAAGGPAQFECETSEAHVHVHVKDGMELGHS 840
Db 781 RGGRIAYQLSVQGLARFLHKDMAGSCYDVAAGGPAQFECETSEAHVHVHVKDGMELGHS 840
Qy 841 GERFLOBDVGTTRHRLVAATVTRQDEGTYSCVGEVSDVDFRLRVSEPKVVFVFAKEOLAREKL 900
Db 841 GERFLOBDVGTTRHRLVAATVTRQDEGTYSCVGEVSDVDFRLRVSEPKVVFVFAKEOLAREKL 900
Qy 901 QAEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVCMEATGCTRRLLVVQAGADAGEY 960
Db 901 QAEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVCMEATGCTRRLLVVQAGADAGEY 960
Qy 961 SCEAGGQRLSFHLDVKEPKVVFVFAKDQVAHSEVQAEAGANATLSCEVAQAQAEVVMYKDGK 1020
Db 961 SCEAGGQRLSFHLDVKEPKVVFVFAKDQVAHSEVQAEAGANATLSCEVAQAQAEVVMYKDGK 1020
Qy 1021 KLSSSLKVHVEAKCRRRLVVQAGKTDAGDYSCEARGQVRSFRLHITPEKMFVAKESQV 1080
Db 1021 KLSSSLKVHVEAKCRRRLVVQAGKTDAGDYSCEARGQVRSFRLHITPEKMFVAKESQV 1080
Qy 1081 HNEVQAEAGASAMLSCEVAQAQTEVTWYKDGKLSSSSKVGMVKGCTRRLLVLPQAGKAD 1140
Db 1081 HNEVQAEAGASAMLSCEVAQAQTEVTWYKDGKLSSSSKVGMVKGCTRRLLVLPQAGKAD 1140
Qy 1141 AGEYSCBAGGQVRSFHLHITPEKGVFAKEQSVHNEVQAEAGTATMLSCVQAQPTQVETWY 1200
Db 1141 AGEYSCBAGGQVRSFHLHITPEKGVFAKEQSVHNEVQAEAGTATMLSCVQAQPTQVETWY 1200
Qy 1201 KDGKLSSSSKVNRVKGCTRRLLVVQVKGADAGEYSCVAGQVRSFOLHITPEKAVFAK 1260
Db 1201 KDGKLSSSSKVNRVKGCTRRLLVVQVKGADAGEYSCVAGQVRSFOLHITPEKAVFAK 1260
Qy 1261 EQLVHNEVTRTEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVRIEAGCWRQLVVOQA 1320
Db 1261 EQLVHNEVTRTEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVRIEAGCWRQLVVOQA 1320
Qy 1321 GQADAGBYTCBAGQRLSFHLDVSEPKAVFAKEQLAHRKVQAEAGATATLSCEVAQAQTE 1380
Db 1321 GQADAGBYTCBAGQRLSFHLDVSEPKAVFAKEQLAHRKVQAEAGATATLSCEVAQAQTE 1380
Qy 1381 VTWYKDGKLSSSSKVRMEAVGCTRRLLVVQACQADTGEYSCBAGGQRLSFSLDVAEPKV 1440
Db 1381 VTWYKDGKLSSSSKVRMEAVGCTRRLLVVQACQADTGEYSCBAGGQRLSFSLDVAEPKV 1440
Qy 1441 VFAKEQPVHREVOQAAGATTTLSCEVAQAQTEVVMYKDGKLSFSSSKVRMEAVGCTRRLLV 1500
Db 1441 VFAKEQPVHREVOQAAGATTTLSCEVAQAQTEVVMYKDGKLSFSSSKVRMEAVGCTRRLLV 1500

QY 1501 VQAGQADAGEYSCAGSORLSFHLHVAEPKAVFAKEQAPASREVQAEAGTSATLSCEVAQ 1560
DB 1501 VQAGQADAGEYSCAGSORLSFHLHVAEPKAVFAKEQAPASREVQAEAGTSATLSCEVAQ 1560
QY 1561 AQTEVTWYKDGKLLSSSKVRMEAVGCTRRLLVYVQAGQADAGEYSCKAGDQRLSFLHJVA 1620
DB 1561 AQTEVTWYKDGKLLSSSKVRMEAVGCTRRLLVYVQAGQADAGEYSCKAGDQRLSFLHJVA 1620
QY 1621 EPKVVPFAKEOPAHREVQAEAGASATLSCEVAQAQTEVTWYKDGKLLSSSKVRVBAVGCT 1680
DB 1621 EPKVVPFAKEOPAHREVQAEAGASATLSCEVAQAQTEVTWYKDGKLLSSSKVRVBAVGCT 1680
QY 1681 RRLVYVQAGQADAGEYSCAGSORLSFHLHVAELEPOJSERPCREPLVVKHEHDIILTA 1740
DB 1681 RRLVYVQAGQADAGEYSCAGSORLSFHLHVAELEPOJSERPCREPLVVKHEHDIILTA 1740
QY 1741 TLATPSAATVTLKDGVEIRRSKRHETASQGDTHLTIVHGAQVLDLSAIYSCRVAEGQDF 1800
DB 1741 TLATPSAATVTLKDGVEIRRSKRHETASQGDTHLTIVHGAQVLDLSAIYSCRVAEGQDF 1800
QY 1801 PVQVEEVAAKFCRLLEPVCGBELGGTVTLACELSPACAEVVMRCGNTQPRVGRFQWVAEG 1860
DB 1801 PVQVEEVAAKFCRLLEPVCGBELGGTVTLACELSPACAEVVMRCGNTQPRVGRFQWVAEG 1860
QY 1861 PVRSLTIVLGLRAEDAGEYVCESRDDHTSAQLTVSVPRVVKFMSGJSTVVAEBGGATFQC 1920
DB 1861 PVRSLTIVLGLRAEDAGEYVCESRDDHTSAQLTVSVPRVVKFMSGJSTVVAEBGGATFQC 1920
QY 1921 VVPSDVAVMFRDQALLQPSKFAISQSGASHSLTISDLVLEDAGQITVBAEGASSAA 1980
DB 1921 VVPSDVAVMFRDQALLQPSKFAISQSGASHSLTISDLVLEDAGQITVBAEGASSAA 1980
QY 1981 LVRREAPVLFKKLEPQTVBERSSVTLVELTRPWPPELRWTRNATAPALAPGKNVEIHBAGA 2040
DB 1981 LVRREAPVLFKKLEPQTVBERSSVTLVELTRPWPPELRWTRNATAPALAPGKNVEIHBAGA 2040
QY 2041 RHRLVLHNVGFADRGFFCETPDDTKOAKLTJMEMQVRLVRGLQAVEAREOGTATMEVOL 2100
DB 2041 RHRLVLHNVGFADRGFFCETPDDTKOAKLTJMEMQVRLVRGLQAVEAREOGTATMEVOL 2100
QY 2101 SHADVDGSWTRDGLRFQOQPTCHLAVRGPMTHTLTLGLRPEDSGJLWVFKAEVHTSARLV 2160
DB 2101 SHADVDGSWTRDGLRFQOQPTCHLAVRGPMTHTLTLGLRPEDSGJLWVFKAEVHTSARLV 2160
QY 2161 VTELVPVSPRPLQDVVTTKEKVTLECELSRPNVDVRLKGVLEUAGKTMAIAAQACR 2220
DB 2161 VTELVPVSPRPLQDVVTTKEKVTLECELSRPNVDVRLKGVLEUAGKTMAIAAQACR 2220
QY 2221 SLTIYRCBFADQGVVCDADHAQSSASVKVQGRVTYTLIYRRLVAEDAGEIQFVAENAESR 2280
DB 2221 SLTIYRCBFADQGVVCDADHAQSSASVKVQGRVTYTLIYRRLVAEDAGEIQFVAENAESR 2280
QY 2281 AQLRVKELPVTILVRPLDKIAHEKHGVLCEQVSRASQVRWFKGSQELQPGPKYELVSD 2340
DB 2281 AQLRVKELPVTILVRPLDKIAHEKHGVLCEQVSRASQVRWFKGSQELQPGPKYELVSD 2340
QY 2341 GLYRKLIIISDVHAEDEDTYTCADGVKTSQAQFVBEQSIITIVRGLODVTVMPEPAPWPEC 2400
DB 2341 GLYRKLIIISDVHAEDEDTYTCADGVKTSQAQFVBEQSIITIVRGLODVTVMPEPAPWPEC 2400
QY 2401 ETSIPSVRPPKWLKGKTVLQAGNVGLQEBCGTVHRLMLRRTCTMTGPHFTVVGSRSSA 2460
DB 2401 ETSIPSVRPPKWLKGKTVLQAGNVGLQEBCGTVHRLMLRRTCTMTGPHFTVVGSRSSA 2460
QY 2461 RLWVSDIPVLTPLPEPTKGRELSQVLSLSCDPRPAPKAVQVYKDDTPLSPSEKFKMSLEG 2520
DB 2461 RLWVSDIPVLTPLPEPTKGRELSQVLSLSCDPRPAPKAVQVYKDDTPLSPSEKFKMSLEG 2520
QY 2521 QMAELRIILRLMPADAGVYRCQAGSAHSSTEVTVAREVTVTGPLQDAEATBEGWASPSCE 2580
DB 2521 QMAELRIILRLMPADAGVYRCQAGSAHSSTEVTVAREVTVTGPLQDAEATBEGWASPSCE 2580
QY 2581 LSHDEEEVWNSLNGHPLYNDSFHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTSARL 2640

DB 2581 LSHDEEEVWNSLNGHPLYNDSFHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTSARL 2640
QY 2641 EVRVKPVVFLKALDDLSAEERGTLALQCEVSDPEAHVVMRKDGVQLGSPDKYDFLHTAGT 2700
DB 2641 EVRVKPVVFLKALDDLSAEERGTLALQCEVSDPEAHVVMRKDGVQLGSPDKYDFLHTAGT 2700
QY 2701 RGLVVDHVSYPEDAGLYTCHVGSSETRARVRVHDLHVGITKRLKTMVEVLEGBSCSCECVLS 2760
DB 2701 RGLVVDHVSYPEDAGLYTCHVGSSETRARVRVHDLHVGITKRLKTMVEVLEGBSCSCECVLS 2760
QY 2761 HESASDPAMWTVGGKTVGSSSRFOATROGRKYILVVRREAAAPSDAGEVVSVRGLTSKASL 2820
DB 2761 HESASDPAMWTVGGKTVGSSSRFOATROGRKYILVVRREAAAPSDAGEVVSVRGLTSKASL 2820
QY 2821 IVRERPAAIIPLEDQWVAPGEDVELRCESLRACTPVHMLKDRKAIRKSQKYDVVCEGTM 2880
DB 2821 IVRERPAAIIPLEDQWVAPGEDVELRCESLRACTPVHMLKDRKAIRKSQKYDVVCEGTM 2880
QY 2881 AMLVIRGASLKDAGEYTCVEASASTASLHVEEKANCFTBELTNLOVEEKGTAFTVCKTE 2940
DB 2881 AMLVIRGASLKDAGEYTCVEASASTASLHVEEKANCFTBELTNLOVEEKGTAFTVCKTE 2940
QY 2941 HPAATVTVWRKGLLELRASGKHQPSQEGTLRLTISALEKADSDTYTCDIGQASQAQLLV 3000
DB 2941 HPAATVTVWRKGLLELRASGKHQPSQEGTLRLTISALEKADSDTYTCDIGQASQAQLLV 3000
QY 3001 QGRRVHIIILEDVVDVQEGSSATFRCRISPANYEPVHMFLLDKPTLHANELNEIDAQPGGY 3060
DB 3001 QGRRVHIIILEDVVDVQEGSSATFRCRISPANYEPVHMFLLDKPTLHANELNEIDAQPGGY 3060
QY 3061 HVLTLRQALXKDSGTIYFEAGDQASAAALRVTERKPSVFSRELTDATITTEGEDTLVLCETS 3120
DB 3061 HVLTLRQALXKDSGTIYFEAGDQASAAALRVTERKPSVFSRELTDATITTEGEDTLVLCETS 3120
QY 3121 TCDIPMCWTOKGKTLRGSARCOLSHEGHRALQLLTGATLQDSGRYKCBAGGACSSSIVRV 3180
DB 3121 TCDIPMCWTOKGKTLRGSARCOLSHEGHRALQLLTGATLQDSGRYKCBAGGACSSSIVRV 3180
QY 3181 HARVPRFOBALKDLLEVLEGGAAATLRCVLSVAAPVKWCYGNVLPAGDKYSLRQSGAMLE 3240
DB 3181 HARVPRFOBALKDLLEVLEGGAAATLRCVLSVAAPVKWCYGNVLPAGDKYSLRQSGAMLE 3240
QY 3241 LVVNRNLRPDQSGRYSCSFGDQTTSATLTVTALPAQFICKLRNKEATEGATATLRCELSKT 3300
DB 3241 LVVNRNLRPDQSGRYSCSFGDQTTSATLTVTALPAQFICKLRNKEATEGATATLRCELSKT 3300
QY 3301 APVEMRGSETLRDGDYCLRDQDGMCELQIRGLAWDAAAEYSCVCGBERTSASLTIIRPM 3360
DB 3301 APVEMRGSETLRDGDYCLRDQDGMCELQIRGLAWDAAAEYSCVCGBERTSASLTIIRPM 3360
QY 3361 PAHFIGRLRHQESSIEGATATLRCELSKAAPVEMRGRESLRDGRHSLRQDGAVCCELOQC 3420
DB 3361 PAHFIGRLRHQESSIEGATATLRCELSKAAPVEMRGRESLRDGRHSLRQDGAVCCELOQC 3420
QY 3421 GLAVADAGEYSCVCGBERTSATLTIVKALPAKFTTEGLRNEEAVEGATAMLWELSKVAPVE 3480
DB 3421 GLAVADAGEYSCVCGBERTSATLTIVKALPAKFTTEGLRNEEAVEGATAMLWELSKVAPVE 3480
QY 3481 WRKGPENLRDGDYTLROEGTRCELOICGLAMADAGEYLCVCGBERTSATLTIRALPARF 3540
DB 3481 WRKGPENLRDGDYTLROEGTRCELOICGLAMADAGEYLCVCGBERTSATLTIRALPARF 3540
QY 3541 IEDVKNQREAREGATAVLQCELSNAAPVEMRGSETLRDGRYSLRQDGTKCELOIRGLAM 3600
DB 3541 IEDVKNQREAREGATAVLQCELSNAAPVEMRGSETLRDGRYSLRQDGTKCELOIRGLAM 3600
QY 3601 ADTGEYSVCVCGBERTSAMLTVRALPIKFTTEGLRNEEATEGATAVLRCBLSKMAPVEMWKG 3660
DB 3601 ADTGEYSVCVCGBERTSAMLTVRALPIKFTTEGLRNEEATEGATAVLRCBLSKMAPVEMWKG 3660
QY 3661 HETLRDGDGRHSLRDQDGMCELQIRGLVAEDAGEYLCMCKERTSAMLTVRAMPSPFIEGL 3720

Db 3661 HETLRDGRHSURQDGCARCELOIRGLVABDAGEYLCMCGKERTSAMLTVRAMPSKFI EBL 3720
Qy 3721 RNEEATEGDTATLWCELSKAAAPVEMRKGHETLRDGRHSRLRQDGRKCELOIRGLAVVDAG 3780
Db 3721 RNEEATEGDTATLWCELSKAAAPVEMRKGHETLRDGRHSRLRQDGRKCELOIRGLAVVDAG 3780
Qy 3781 EYSCVCGQERTSATLTVRALPARFIEDVKNQBARAGATAVLQCELSKAAPVEMRKGS ETL 3840
Db 3781 EYSCVCGQERTSATLTVRALPARFIEDVKNQBARAGATAVLQCELSKAAPVEMRKGS ETL 3840
Qy 3841 RCGDRYSLRQDGTRELOIHLGLSVADTGEYSVCVCGQERTSATLTVRALPARFIEDVKNQBARAGATAVLQCELSKAAPVEMRKGS ETL 3900
Db 3841 RCGDRYSLRQDGTRELOIHLGLSVADTGEYSVCVCGQERTSATLTVRALPARFIEDVKNQBARAGATAVLQCELSKAAPVEMRKGS ETL 3900
Qy 3901 ABEESTATLQCELSBPTATVWWSKGLQLOANGREPRLOQCTAELVLDIQLQREDTGEYT 3960
Db 3901 ABEESTATLQCELSBPTATVWWSKGLQLOANGREPRLOQCTAELVLDIQLQREDTGEYT 3960
Qy 3961 CTCGSOATSATLTVTAAPVRFURELQHOEVDGEGTAHLCCELSRAGASVEMRKGSLOLFP 4020
Db 3961 CTCGSOATSATLTVTAAPVRFURELQHOEVDGEGTAHLCCELSRAGASVEMRKGSLOLFP 4020
Qy 4021 CAKYQWQDGAABEALLVRGVEQEDAGDYTCDTGHTQSHASLSVRPRPKFKTRLOSLEQ E 4080
Db 4021 CAKYQWQDGAABEALLVRGVEQEDAGDYTCDTGHTQSHASLSVRPRPKFKTRLOSLEQ E 4080
Qy 4081 TGDIALCCOLSDAESGAVVQWLKEGVHLAGPKYEMRSQGAATRELLIHOLEAKDTGEYA 4140
Db 4081 TGDIALCCOLSDAESGAVVQWLKEGVHLAGPKYEMRSQGAATRELLIHOLEAKDTGEYA 4140
Qy 4141 CVTGQOKTAASLRVTEPEVTIVRGLVDAEVTADEVDFECSVSRAGATGVQWCLQGLPLQ 4200
Db 4141 CVTGQOKTAASLRVTEPEVTIVRGLVDAEVTADEVDFECSVSRAGATGVQWCLQGLPLQ 4200
Qy 4201 SNEVTEVAVRDRGRIHTLRLKGVTPEDAGTVSPHLGNHASSAQLTVRAPEVTILEPLQDVQ 4260
Db 4201 SNEVTEVAVRDRGRIHTLRLKGVTPEDAGTVSPHLGNHASSAQLTVRAPEVTILEPLQDVQ 4260
Qy 4261 LSEGDQASFOCRLSRASQOEARWALGGVPLQANEMNDITVBOGTLHLHLTLHKVTL EDA GT 4320
Db 4261 LSEGDQASFOCRLSRASQOEARWALGGVPLQANEMNDITVBOGTLHLHLTLHKVTL EDA GT 4320
Qy 4321 VSFHVGTCSBAQLKVTAKNTVVRGLENVLEALEGCEALFECOLSQOPEVAATHWLLDD E PV 4380
Db 4321 VSFHVGTCSBAQLKVTAKNTVVRGLENVLEALEGCEALFECOLSQOPEVAATHWLLDD E PV 4380
Qy 4381 RTSENAENVFPENGIRHLLLLKNLRPQDS CRVTFLAGDMVTSAFITVRGWRLEI L E PLKN 4440
Db 4381 RTSENAENVFPENGIRHLLLLKNLRPQDS CRVTFLAGDMVTSAFITVRGWRLEI L E PLKN 4440
Qy 4441 AAVRAGAQAARFTCTLSEAVPVGEASWYINGAAVQPDSDWTATADGSHQALLLSAOPHH 4500
Db 4441 AAVRAGAQAARFTCTLSEAVPVGEASWYINGAAVQPDSDWTATADGSHQALLLSAOPHH 4500
Qy 4501 AGEVTFACRDAVASARLTVGLPDPEDAEEVVAHSSHTVTLSWAAPMSDGGGCGYRVE 4560
Db 4501 AGEVTFACRDAVASARLTVGLPDPEDAEEVVAHSSHTVTLSWAAPMSDGGGCGYRVE 4560
Qy 4561 VKEGATGQWRLCHELVPCEVCVVDGLAPGETYRFRVAAGVPGAGEPVHLPTQVRLAEP P 4620
Db 4561 VKEGATGQWRLCHELVPCEVCVVDGLAPGETYRFRVAAGVPGAGEPVHLPTQVRLAEP P 4620
Qy 4621 KPVPQPSAPESRQVAAGEDVLSLEVVAAEAGEVIWHKGMERIQPGGRFEVVUSQGRQOML 4680
Db 4621 KPVPQPSAPESRQVAAGEDVLSLEVVAAEAGEVIWHKGMERIQPGGRFEVVUSQGRQOML 4680
Qy 4681 VIKGFTABDQGEYHCGLAQGSICPAAATFQVALSPASVDEAPQPSLPPEAAQEGDLHL L W 4740
Db 4681 VIKGFTABDQGEYHCGLAQGSICPAAATFQVALSPASVDEAPQPSLPPEAAQEGDLHL L W 4740
Qy 4741 EALARKRMSREPTLDSISELPEEDGRSQRLOPEAREEVAAPDLSEGYSTADELATGDA D L 4800
Db 4741 EALARKRMSREPTLDSISELPEEDGRSQRLOPEAREEVAAPDLSEGYSTADELATGDA D L 4800

Qy 4801 SHTSDDDESRACTPSLVTYLLKAGRPCTSPLASKVGAPAAAPSVKPOQOQEEPLAAVRPPLG 4860
Db 4801 SHTSDDDESRACTPSLVTYLLKAGRPCTSPLASKVGAPAAAPSVKPOQOQEEPLAAVRPPLG 4860
Qy 4861 DLSTKOLGDPMDKAAAVKIQAAAFKGYKVRKEMKQOEGPMFSHTFGDTBAQVGDALRL E CV 4920
Db 4861 DLSTKOLGDPMDKAAAVKIQAAAFKGYKVRKEMKQOEGPMFSHTFGDTBAQVGDALRL E CV 4920
Qy 4921 VASKADYRARWLKDGVELTDGRHHHIDQLDGGTCSLLIAGLDRADACGYTCOVSNKFGO V 4980
Db 4921 VASKADYRARWLKDGVELTDGRHHHIDQLDGGTCSLLIAGLDRADACGYTCOVSNKFGO V 4980
Qy 4981 THSACVVVSGSESAESSGGELDDAFRAARRLHRLFRTKSPAESVDEELFLSADEGPA 5040
Db 4981 THSACVVVSGSESAESSGGELDDAFRAARRLHRLFRTKSPAESVDEELFLSADEGPA 5040
Qy 5041 EPEEPADQWTVREDEHFCIRFEALTEARQAVTRFQEMFATLGI GVEIKLVEQGPRR VEM 5100
Db 5041 EPEEPADQWTVREDEHFCIRFEALTEARQAVTRFQEMFATLGI GVEIKLVEQGPRR VEM 5100
Qy 5101 CISKETPAPVVPPEPLSLTSDAAPVFLTELQOEVDQGYPVSPDCVVTGQPMPSVRWF 5160
Db 5101 CISKETPAPVVPPEPLSLTSDAAPVFLTELQOEVDQGYPVSPDCVVTGQPMPSVRWF 5160
Qy 5161 KDGKLLBEDDHYMINEQQQGHQLIITAVVPADMGVYRCLAENSMGVSSTKAE LRVDLTS 5220
Db 5161 KDGKLLBEDDHYMINEQQQGHQLIITAVVPADMGVYRCLAENSMGVSSTKAE LRVDLTS 5220
Qy 5221 TDYTAADATESSSYFSAQGYLSRREOEGTSTTDEGOLPOVVEBELRDLQVAPGTR LAKF 5280
Db 5221 TDYTAADATESSSYFSAQGYLSRREOEGTSTTDEGOLPOVVEBELRDLQVAPGTR LAKF 5280
Qy 5281 QLKVKGYPA RLYWFKDGOPLTASAHIRMTDKKILHTLEIISVTR EDSOGVAA YISNMG 5340
Db 5281 QLKVKGYPA RLYWFKDGOPLTASAHIRMTDKKILHTLEIISVTR EDSOGVAA YISNMG 5340
Qy 5341 AAYSARLLVRGPDEPEKEPASDVHEQLVPPMLERFTPKKVKKGSSITFSVKVEGR PVP 5400
Db 5341 AAYSARLLVRGPDEPEKEPASDVHEQLVPPMLERFTPKKVKKGSSITFSVKVEGR PVP 5400
Qy 5401 TVHMLREABRGVUNI GPDTPGYTVASSAOHSILVLDVGRHQHGTTCI TASNAAGQALC 5460
Db 5401 TVHMLREABRGVUNI GPDTPGYTVASSAOHSILVLDVGRHQHGTTCI TASNAAGQALC 5460
Qy 5461 SASLHVSGLPKVEBOEKKEALISTFLOGTTQATISAOGLETTASADLGQKKEP LAAKE 5520
Db 5461 SASLHVSGLPKVEBOEKKEALISTFLOGTTQATISAOGLETTASADLGQKKEP LAAKE 5520
Qy 5521 ALGHLSLAEVGTBEFLQKLTQSQITEMVSAKITQAKLVPGGSDSDSKTSPASPRHGRSR 5580
Db 5521 ALGHLSLAEVGTBEFLQKLTQSQITEMVSAKITQAKLVPGGSDSDSKTSPASPRHGRSR 5580
Qy 5581 PSSSIQSSSESDGDARGEI FDIYVVTADYLPUGASODAITLREGQVVEVLDAAHPLR W 5640
Db 5581 PSSSIQSSSESDGDARGEI FDIYVVTADYLPUGASODAITLREGQVVEVLDAAHPLR W 5640
Qy 5641 LVRTKPTKSSPSRQGWSPAYLDRRLKLSPEWGAABEPPEPGEAVSDEYKARLS SSVIOE 5700
Db 5641 LVRTKPTKSSPSRQGWSPAYLDRRLKLSPEWGAABEPPEPGEAVSDEYKARLS SSVIOE 5700
Qy 5701 LLSSEQA FVBELOP LQSHHLLQHLERCPHPVIAVAGQKAVIFRNVRD IGRPHSSFLQELQ Q 5760
Db 5701 LLSSEQA FVBELOP LQSHHLLQHLERCPHPVIAVAGQKAVIFRNVRD IGRPHSSFLQELQ Q 5760
Qy 5761 CDTDDDDVAMCFIKNOAAFEQYLEFLVGRVOAESVWVSTAIQEFYKVKYABEALLAGDPSQ P 5820
Db 5761 CDTDDDDVAMCFIKNOAAFEQYLEFLVGRVOAESVWVSTAIQEFYKVKYABEALLAGDPSQ P 5820
Qy 5821 PPPPLQHYLQOPVERVORYQALLKELRNKARNQNOCALLEQA VAVVSALLPQRAENKLHV 5880
Db 5821 PPPPLQHYLQOPVERVORYQALLKELRNKARNQNOCALLEQA VAVVSALLPQRAENKLHV 5880

QY 5881 SLWENY PGTLEALGPPIROGHFIVMEGAPCAMPKGHNHVFLEFRNLHVICKPRDSRT 5940
 Db 5881 SLWENY PGTLEALGPPIROGHFIVMEGAPCAMPKGHNHVFLEFRNLHVICKPRDSRT 5940
 QY 5941 DTVSVFRNMKLSIDLDQVEGDDRAFEVQWEREDSVRKYLLOARTAIKSSWKEIC 6000
 Db 5941 DTVSVFRNMKLSIDLDQVEGDDRAFEVQWEREDSVRKYLLOARTAIKSSWKEIC 6000
 QY 6001 GIQORLALPVWRPPPEBELADCTAELGETVKLACRVGTGTPKPVISWYKDGKAVQVDPHH 6060
 Db 6001 GIQORLALPVWRPPPEBELADCTAELGETVKLACRVGTGTPKPVISWYKDGKAVQVDPHH 6060
 QY 6061 ILIEDPDGSCALILDSLTGVDSSQYMCFAASAGNCSTLGLKVQVPRFVNKVRASPPV 6120
 Db 6061 ILIEDPDGSCALILDSLTGVDSSQYMCFAASAGNCSTLGLKVQVPRFVNKVRASPPV 6120
 QY 6121 EGEDAQFTCTIEGAPYQIRWYKDGALLTGNKFTLSEPRSGLLVLVIRAAKEDLGLY 6180
 Db 6121 EGEDAQFTCTIEGAPYQIRWYKDGALLTGNKFTLSEPRSGLLVLVIRAAKEDLGLY 6180
 QY 6181 ECELVNRLGSARASAEIRIQSPMLQAQEQCHREQLVAIVEDTTLER-----ADQEV 6232
 Db 6181 ECELVNRLGSARASAEIRIQSPMLQAQEQCHREQLVAIVEDTTLER-----ADQEV 6232
 QY 6233 SVLKRLGLPKAPGPSTGDLTGPGCPRG-----APAL-----QETGSP- 6271
 Db 6241 TVVKSPPROQR--SPKSPSPSPSCASPLRPLGALLDLYLPGAGQRRPRAEPGQKPV 6299
 QY 6272 -PVTGTSEAPVPPRV-----PQPLLHEGPE-----QEPEAIARAQEQWTPIRMEGA 6317
 Db 6300 VPTLVVTEAEAHSPALPGLSGPQPKWVEEETIEVRVKMGPGQVSPTE--VPRSSGH 6357
 QY 6318 AW--PQAGTG 6325
 Db 6358 LFTLPGATPG 6367

RESULT 3

ABG76186
 ID ABG76186 standard; protein; 2630 AA.
 XX
 AC ABG76186;
 DT 09-MAY-2003 (first entry)
 XX
 DE Human serine/threonine or protein kinase 59079.
 KW Human; enzyme; serine/threonine kinase; protein kinase; 59079;
 KW cardiovascular disease; heart failure; myocardial infarction;
 KW blood vessel disorder; atherosclerosis; Kaposi's sarcoma; immunogen;
 KW blood platelet disorder; thrombocytopaenia; leukaemia; Hodgkin's disease;
 KW haemolytic anaemia; cellular proliferative disorder; cancer;
 KW protein kinase disorder; autoimmune disorder; diabetes mellitus;
 KW psoriasis; inflammatory bowel disease; rheumatoid arthritis;
 KW multiple sclerosis.
 XX
 OS Homo sapiens.
 XX
 PN US2002168742-A1.
 XX
 PD 14-NOV-2002.
 XX
 PF 15-FEB-2002; 2002US-00077130.
 XX
 PR 15-FEB-2001; 2001US-0269201P.
 XX
 XX (MILL-) MILLENNIUM PHARM INC.
 PA Kapeller-Libermann R, Acton SL;
 PI WPI; 2003-298729/29.
 XX
 DR N-PSDB; ABX11641.
 XX

PT Novel isolated human protein kinase, designated 59079 or 12599
 PT polypeptide, useful as diagnostic and therapeutic agents for preventing
 PT cardiovascular diseases, proliferative disorders, and protein kinase
 PT disorders.
 XX Claim 8; Page 48-54; 119pp; English.
 XX
 CC The invention relates to an isolated human serine/threonine or protein
 CC kinase, 59079 or 12599 polypeptide, encoded by nucleic acid molecule
 CC comprising at least 85% identity to the nucleic acids appearing as
 CC ABX11641 and ABX11642 or their complement, a naturally occurring variant
 CC of the kinases or their fragments. Also included are a non-human host
 CC cell containing the nucleic acids, an antibody specific for the proteins,
 CC identifying a compound which binds to the kinase (by contacting the
 CC kinase or a cell expressing the kinase with a test compound and
 CC determining whether the kinase binds to the test compound) and modulating
 CC the activity of kinase using the identified compound. The kinases and
 CC their encoding nucleic acids are useful as diagnostic and therapeutic
 CC agents for preventing a disease or condition associated with an aberrant
 CC or unwanted 59079 or 12599 activity in a subject, including
 CC cardiovascular diseases such as heart failure, and myocardial infarction;
 CC disorders involving blood vessels such as atherosclerosis, and Kaposi's
 CC sarcoma; blood platelets disorder such as thrombocytopaenia, leukaemia,
 CC Hodgkin's disease, haemolytic anaemia; cellular proliferative disorders
 CC such as cancer; and protein kinase disorders such as autoimmune
 CC disorders, diabetes mellitus, psoriasis, inflammatory bowel disease,
 CC rheumatoid arthritis, and multiple sclerosis (many examples of diseases
 CC and disorders are included in the specification). The kinases, their
 CC encoding nucleic acids and antibodies are useful in screening assays,
 CC detection assays (e.g. forensic biology), and predictive medicine (e.g.
 CC diagnostic assays, prognostic assays, and monitoring clinical trials and
 CC pharmacogenomics). The kinases and their encoding nucleic acids are
 CC useful as query sequences to perform a search against public databases to
 CC identify other family members or related sequences. The present sequence
 CC represents kinase 59079
 XX
 SQ Sequence 2630 AA;
 Query Match 33.2%; Score 13710; DB 6; Length 2630;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 5339 MGAAYSSARLLVRGDPDEPEEKPSADVHEQLVPPRMLEFPTPKVKKSSITFSVKEGRP 5398
 Db 1 MGAAYSSARLLVRGDPDEPEEKPSADVHEQLVPPRMLEFPTPKVKKSSITFSVKEGRP 60
 QY 5399 VPTVHMLREEAERGVLWIGPDTPGYTVASSAQHSLVLLDVGRHQGTTCIASNAAGQA 5458
 Db 61 VPTVHMLREEAERGVLWIGPDTPGYTVASSAQHSLVLLDVGRHQGTTCIASNAAGQA 120
 QY 5459 LCSASLHVSGLPKVEEQEKVKEALISTFLOGTQTAISAGLETAFADLGQRKKEPLAA 5518
 Db 121 LCSASLHVSGLPKVEEQEKVKEALISTFLOGTQTAISAGLETAFADLGQRKKEPLAA 180
 QY 5519 KENLGHLSLAEVGTTEFLQKLTSTQITEMVSAKITQAKLVQPGDSDSDSKTSASPRHGR 5578
 Db 181 KENLGHLSLAEVGTTEFLQKLTSTQITEMVSAKITQAKLVQPGDSDSDSKTSASPRHGR 240
 QY 5579 SRPSSSIOESSSESDGDARGEIFDIYVVTADYLPGLAEQDAITLREGYVVLDAHPL 5638
 Db 241 SRPSSSIOESSSESDGDARGEIFDIYVVTADYLPGLAEQDAITLREGYVVLDAHPL 300
 QY 5639 RNLVTRTKTSPSRQGWVSPAYLDRRLKLSPEWGAABEAPFPGSAVSEDEYKARLSSVI 5698
 Db 301 RNLVTRTKTSPSRQGWVSPAYLDRRLKLSPEWGAABEAPFPGSAVSEDEYKARLSSVI 360
 QY 5699 QELLSSEQA FVEELQFLOSHLQHLERCPHPVIAVAGQKAVIFRNVRDIGRPHSSFLQEL 5758
 Db 361 QELLSSEQA FVEELQFLOSHLQHLERCPHPVIAVAGQKAVIFRNVRDIGRPHSSFLQEL 420
 QY 5759 QOCOTDDDDVAMCFIKNOAAFEQYLEFLVGRVQAESVWVSTAIQEFYKKYABEALLAGDPS 5818
 Db 421 QOCOTDDDDVAMCFIKNOAAFEQYLEFLVGRVQAESVWVSTAIQEFYKKYABEALLAGDPS 480

Qy	5819	QPPPPPIQHYLEQPVVERVQRYOALLKELIRNKARNRQNCALLBOAYAVVSGALLFORAENKL	5878
Db	481	QPPPPPIQHYLEQPVVERVQRYOALLKELIRNKARNRQNCALLBOAYAVVSGALLFORAENKL	540
Qy	5879	HVSLWMENYPGTLEALGBPIROGFHIVWEGAPGAMPKGNHRHVFLFRNHLVICKPRRDS	5938
Db	541	HVSLWMENYPGTLEALGBPIROGFHIVWEGAPGAMPKGNHRHVFLFRNHLVICKPRRDS	600
Qy	5939	RTDTSVYVFRNMKLSIDLNDQVEGDDRAFEVMQEREDSVRKYLLQARTAIIKSSWVKE	5998
Db	601	RTDTSVYVFRNMKLSIDLNDQVEGDDRAFEVMQEREDSVRKYLLQARTAIIKSSWVKE	660
Qy	5999	ICGTQOBRALPVMRPPPFEBELADCTAELGETVKLACRVGTGPKPVIWTKGOKAVQVDP	6058
Db	661	ICGTQOBRALPVMRPPPFEBELADCTAELGETVKLACRVGTGPKPVIWTKGOKAVQVDP	720
Qy	6059	HHILIEBPDGSCALILDSLTGVSQGYMCFAASAAGNCSTLGKILVQVPRFVNVKVRASP	6118
Db	721	HHILIEBPDGSCALILDSLTGVSQGYMCFAASAAGNCSTLGKILVQVPRFVNVKVRASP	7800
Qy	6119	FVEGEDAQFTCTIEGAPYPOIRMYKDGALLTGNKFKQTLSEPRSGLLVLVIRAASKEDLG	6178
Db	781	FVEGEDAQFTCTIEGAPYPOIRMYKDGALLTGNKFKQTLSEPRSGLLVLVIRAASKEDLG	840
Qy	6179	LYECELVNRIGLSARASAEILRIQSPMLQAOQCCHREQLVAAVEDTTLERADQEVTSVLKRL	6238
Db	841	LYECELVNRIGLSARASAEILRIQSPMLQAOQCCHREQLVAAVEDTTLERADQEVTSVLKRL	900
Qy	6239	LGPKAPGPGTDLTGPGPCPRGAPALQETGSPVTCGTSEAPVPPRPVPOPLLHGEPEQE	6298
Db	901	LGPKAPGPGTDLTGPGPCPRGAPALQETGSPVTCGTSEAPVPPRPVPOPLLHGEPEQE	960
Qy	6299	PEATARAQEWTPVIRMEGAAMPGAGTGCELLDWDVHSHVVRRETTQRTYTAQIDTHTARP	6358
Db	961	PEATARAQEWTPVIRMEGAAMPGAGTGCELLDWDVHSHVVRRETTQRTYTAQIDTHTARP	1020
Qy	6359	MOVTIEDVQACTGTAQFEALIEGDPOPSVTWTKDSVLVDSTRLSQOQEGTYSVLVRH	6418
Db	1021	MOVTIEDVQACTGTAQFEALIEGDPOPSVTWTKDSVLVDSTRLSQOQEGTYSVLVRH	1080
Qy	6419	VASKDAGVYTCLAQNTGGQVLCXAEILLVLGDNEPDESEKSHRRKULHSFYEVKEEIGRGV	6478
Db	1081	VASKDAGVYTCLAQNTGGQVLCXAEILLVLGDNEPDESEKSHRRKULHSFYEVKEEIGRGV	1140
Qy	6479	FGFVKRVQHKGNKILCAAKFIPLRSRTRAQAYRERDILAAALSHPLVTGLLDQFETRKTLI	6538
Db	1141	FGFVKRVQHKGNKILCAAKFIPLRSRTRAQAYRERDILAAALSHPLVTGLLDQFETRKTLI	1200
Qy	6539	LILIELCSSEELLORLYRKGVVTEAEVKVYIQOLVEGLHYLHSHGVFLHDIKPSNILLMVHP	6598
Db	1201	LILIELCSSEELLORLYRKGVVTEAEVKVYIQOLVEGLHYLHSHGVFLHDIKPSNILLMVHP	1260
Qy	6599	AREDIKICDFGFAQNTIPAELOFSQYGSPEFVSPEIIQQNPVSEASDIWAMGVISYLSLT	6658
Db	1261	AREDIKICDFGFAQNTIPAELOFSQYGSPEFVSPEIIQQNPVSEASDIWAMGVISYLSLT	1320
Qy	6659	CSSPFFAGESDRATLLNVLGRVSNWSPMAAHLSEDAKDFIKATLQAPQARPSAAQCLSH	6718
Db	1321	CSSPFFAGESDRATLLNVLGRVSNWSPMAAHLSEDAKDFIKATLQAPQARPSAAQCLSH	1380
Qy	6719	PWFLKSNPAAEEAHFINTKQLFKLLARSWORSLSMSYKSILVWESIPELLRGPPDPSLGV	6778
Db	1381	PWFLKSNPAAEEAHFINTKQLFKLLARSWORSLSMSYKSILVWESIPELLRGPPDPSLGV	1440
Qy	6779	ARHLCRDTGSSSSSSSSSDNELAPFAAKSLPPSPVTHSPLLHPRGFLRPSASLPEEAEA	6838
Db	1441	ARHLCRDTGSSSSSSSSSDNELAPFAAKSLPPSPVTHSPLLHPRGFLRPSASLPEEAEA	1500
Qy	6839	SERSTEAPAPSPGEGAGPPAAQCGVPRHVSIVRSLSFTHQGESPEHGALAPGSRHRPARR	6898
Db	1501	SERSTEAPAPSPGEGAGPPAAQCGVPRHVSIVRSLSFTHQGESPEHGALAPGSRHRPARR	1560

RESULT 4
AAB30569
ID AAB30569 standard; protein; 2596 AA.
AC AAB30569;
XX
XX
DT 19-MAR-2001 (first entry)
XX
XX A splice variant of a signal transduction polypeptide.
DE
XX
XX Signal transduction; H19G5; kinase; cardiac disease; angina pectoris;
KW congestive heart failure; dilated congestive cardiomyopathy;
KW hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension;
KW mitral valve disease; aortic valve disease; tricuspid valve disease;
KW myocardial infarction; cardiac arrhythmia; arteriosclerosis;
KW atherosclerosis; cardiac tumour; microbial infection; splice variant.
XX
XX Homo sapiens.
XX
XX
FH Key Location/Qualifiers
FT Domain 325..504
FT /note= "guanine nucleotide exchange factor domain"
FT Domain 1094..1351
FT /note= "kinase domain"
FT Domain 2301..2553
FT /note= "kinase domain"
XX
XX WO200063381-A1.
PN
XX
XX 26-OCT-2000.
PD
XX
XX 11-APR-2000; 2000WO-US009488.
PF
XX
XX 16-APR-1999; 99US-0129553P.
PR
XX
XX (SCIO-) SCIOS INC.
PA
XX
XX Zeng W, Stanton L, Kong H;
PI
XX
XX WPI; 2001-007013/01.
DR N-PSDB; AAC62287.
DR
XX
XX Novel h19G5 polypeptides capable of regulating signal transduction and
PT exhibiting kinase activity useful for identifying antibodies to treat
FT cardiac diseases, and additional mediators of signal transduction.
XX
XX
PS Claim 1; Page 68-74; 81pp; English.
XX
XX The present sequence represents a splice variant of human in signal
CC transduction polypeptide. The polypeptide is designated H19G5. The
CC protein is capable of regulating signal transduction and exhibits kinase
CC activity. The H19G5 transcript is expressed in the heart. H19G5
CC polypeptides and polynucleotides are useful for preventing or treating a
CC cardiac disease, such as congestive heart failure, dilated congestive
CC cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy,
CC mitral valve disease, aortic valve disease or tricuspid valve disease,
CC angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary,
CC arterial or renovascular hypertension, arteriosclerosis, atherosclerosis
CC and cardiac tumours in humans. The polypeptide is also useful for
CC detecting the expression of a protein capable of regulating signal
CC transduction or the expression of a protein capable of acting as a donor
CC or acceptor molecule of a phosphate group. The monoclonal antibodies can
CC be used as probes for detecting discrete antigens expressed by tissue or
CC cell samples, and therefore used in humans for localization and
CC monitoring of microbial infection
XX
XX Sequence 2596 AA;
Query Match 32.8%; Score 13528; DB 4; Length 2596;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2595; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX 5373 MLERPTPKVKVKGSSITPSVKVEGRPVPTVHVLREAEARGVWIGDTPGTIVASSAQOH 5432

Db 1 MLERPTPKVKVKGSSITPSVKVEGRPVPTVHVLREAEARGVWIGDTPGTIVASSAQOH 60
QY 5433 SLVLLDVGRQHQGTYYTCIASNAAGQALCSASLHVSGLPKVBEEQKVKALISTFLOGTTQ 5492
Db 61 SLVLLDVGRQHQGTYYTCIASNAAGQALCSASLHVSGLPKVBEEQKVKALISTFLOGTTQ 120
QY 5493 AISAQGLTASPADLGGORKEEPLAAKALGHLSLAEVGTTEFLQKLSQITEMVSAKIT 5552
Db 121 AISAQGLTASPADLGGORKEEPLAAKALGHLSLAEVGTTEFLQKLSQITEMVSAKIT 180
QY 5553 QAKLVPGGDEDESKTPSPRGRSRPSSIOESSSESDGDARGIFDIYVVTADYL 5612
Db 181 QAKLVPGGDEDESKTPSPRGRSRPSSIOESSSESDGDARGIFDIYVVTADYL 240
QY 5613 PLGAEQDAITLREGQYVEVLDAAHPLRWLVRTKTKSSPSRQGWVSPAYLDRRLKLSPEW 5672
Db 241 PLGAEQDAITLREGQYVEVLDAAHPLRWLVRTKTKSSPSRQGWVSPAYLDRRLKLSPEW 300
QY 5673 GAAEAPPEGSAVSEDEYKARLSSVIGELLSSQAFVEELQLSHHLOHLERCHPVDIA 5732
Db 301 GAAEAPPEGSAVSEDEYKARLSSVIGELLSSQAFVEELQLSHHLOHLERCHPVDIA 360
QY 5733 VAGOKAVIFRNVRDIGRPHSSFLOEQCDTDDVAMCFIKNQAAFEQYLFVGRVQAE 5792
Db 361 VAGOKAVIFRNVRDIGRPHSSFLOEQCDTDDVAMCFIKNQAAFEQYLFVGRVQAE 420
QY 5793 SVVYSTAIQEFYKYABEALLAGDPSPPPPPLQHYLEQPVVERVQRYQALLKELIRNKAR 5852
Db 421 SVVYSTAIQEFYKYABEALLAGDPSPPPPPLQHYLEQPVVERVQRYQALLKELIRNKAR 480
QY 5853 NRQNCALLEQAYAVVSALPQBAENKLVSLMENVPTGLEALGEPTRQGHFIWEGAPGAR 5912
Db 481 NRQNCALLEQAYAVVSALPQBAENKLVSLMENVPTGLEALGEPTRQGHFIWEGAPGAR 540
QY 5913 MPWKGHNHVFLEFRNLVICKPRRDSRTDTVSYYFRNMKLSIDLNQVGGDDRAFEVW 5972
Db 541 MPWKGHNHVFLEFRNLVICKPRRDSRTDTVSYYFRNMKLSIDLNQVGGDDRAFEVW 600
QY 5973 QEREDSVRKYLLQARTAIKSSWVKEICGQQRLALPVRPDPFEEELADCTAEIGETVK 6032
Db 601 QEREDSVRKYLLQARTAIKSSWVKEICGQQRLALPVRPDPFEEELADCTAEIGETVK 660
QY 6033 LACRVGTGPKVVISWYKDGKAVQVDPHHILIEDPGSCALLIDSLTGVDGQYMCFAASA 6092
Db 661 LACRVGTGPKVVISWYKDGKAVQVDPHHILIEDPGSCALLIDSLTGVDGQYMCFAASA 720
QY 6093 AGNCSTLGLKILVQVPPFRVNVKVRASPFVEGEDAQTCTIEGAPYQIRWYKDGALLTTGN 6152
Db 721 AGNCSTLGLKILVQVPPFRVNVKVRASPFVEGEDAQTCTIEGAPYQIRWYKDGALLTTGN 780
QY 6153 KFOQLSEPRSGLLVLVIRAAKEDGLGYECBLVNRGLSARASAEIRIQSPMLQAEQOCHR 6212
Db 781 KFOQLSEPRSGLLVLVIRAAKEDGLGYECBLVNRGLSARASAEIRIQSPMLQAEQOCHR 840
QY 6213 EQLVAAVEDTTLEADQSVTSVLKRLGPKAPGSTGDLTGCPGCPGAPALQETGSOPP 6272
Db 841 EQLVAAVEDTTLEADQSVTSVLKRLGPKAPGSTGDLTGCPGCPGAPALQETGSOPP 900
QY 6273 VTGTSEAPVPRVPQPLLHGGPQEPPIARAQEWTPVIRMEGAAMPAGTGELLWDVH 6332
Db 901 VTGTSEAPVPRVPQPLLHGGPQEPPIARAQEWTPVIRMEGAAMPAGTGELLWDVH 960
QY 6333 SHVVRETTQRTYTYQAIDTHTARPSSMQVTIEDVQAQTGGTAQFEAIIIEGDPQPSVTWYK 6392
Db 961 SHVVRETTQRTYTYQAIDTHTARPSSMQVTIEDVQAQTGGTAQFEAIIIEGDPQPSVTWYK 1020
QY 6393 DSVQVLDSTRLSQQOEGTYSVLVRHVASKDAGVYVTCIAQNTGGQVLCABELLVLGGDNE 6452
Db 1021 DSVQVLDSTRLSQQOEGTYSVLVRHVASKDAGVYVTCIAQNTGGQVLCABELLVLGGDNE 1080
QY 6453 PDSEKQSHRRKLVHSGFYEVKEEIGRGVFGVQVHKGKNIKCAAKFIPILRSRTRAQAVRE 6512

Db 1081 PDSEKQSHRRKLSHFYEVKEEIGRGVGFVKRVQHKGKNKILCAAKFIPLRSRTRAQAYRE 1140
Qy 6513 RDLAALSHPLVTGLDQDPETKTKTILILELCSSEELLDRLYRKGVTEAEVKVYIQQLV 6572
Db 1141 RDLAALSHPLVTGLDQDPETKTKTILILELCSSEELLDRLYRKGVTEAEVKVYIQQLV 1200
Qy 6573 EGLHYLHSHGVHLHDIKPSNILMVHPAREDIKICDFGFAQNTTPAELQFSQYSGSEFVSP 6632
Db 1201 EGLHYLHSHGVHLHDIKPSNILMVHPAREDIKICDFGFAQNTTPAELQFSQYSGSEFVSP 1260
Qy 6633 EIIQONPVSEASDIWAMGVISYLSLTCSPFAGESDRATLLNVLEGRVSWSSPMAHLSE 6692
Db 1261 EIIQONPVSEASDIWAMGVISYLSLTCSPFAGESDRATLLNVLEGRVSWSSPMAHLSE 1320
Qy 6693 DAKDFIKATLQAPQAPSAACLSHPWFLKMPAEAEAHFINTKQLKFLIARSRQORSIM 6752
Db 1321 DAKDFIKATLQAPQAPSAACLSHPWFLKMPAEAEAHFINTKQLKFLIARSRQORSIM 1380
Qy 6753 SYKSILVMRSIPPELLRGPPDSPSLGVARHLCRDTGGSSSSSSSDNELAPFARAKSLPPS 6812
Db 1381 SYKSILVMRSIPPELLRGPPDSPSLGVARHLCRDTGGSSSSSSSDNELAPFARAKSLPPS 1440
Qy 6813 PYTHSPLHPRGFLRPSASLPEAEASERSTAPAPPASPEGAGPAAQGCVPRHVSIRS 6872
Db 1441 PYTHSPLHPRGFLRPSASLPEAEASERSTAPAPPASPEGAGPAAQGCVPRHVSIRS 1500
Qy 6873 LFYHQAGSPHEGALAPGRRHPRARRHLLKGGYIAGALPGLREPLMEHRVLEEAAREE 6932
Db 1501 LFYHQAGSPHEGALAPGRRHPRARRHLLKGGYIAGALPGLREPLMEHRVLEEAAREE 1560
Qy 6933 QATLLAKAPSFETALURLPASGTHLAPGSHSLSHDPSSTPRPSSACGEAQLRPSAPSGG 6992
Db 1561 QATLLAKAPSFETALURLPASGTHLAPGSHSLSHDPSSTPRPSSACGEAQLRPSAPSGG 1620
Qy 6993 APIRDMGHPOGSKQLPSTGGCHPTAQPERPSPDSFWGQPAFCHPKQGSAPQEGCSHPA 7052
Db 1621 APIRDMGHPOGSKQLPSTGGCHPTAQPERPSPDSFWGQPAFCHPKQGSAPQEGCSHPA 1680
Qy 7053 VAPCPGSGFPFGSCKEAPLVSSPPLGQOPAPPAPAKASPLDLSKWGPGDISLPGRPKPG 7112
Db 1681 VAPCPGSGFPFGSCKEAPLVSSPPLGQOPAPPAPAKASPLDLSKWGPGDISLPGRPKPG 1740
Qy 7113 PCSSPFGSASQASSQVSSLRVSSQVGTGTEPGSLDAGWTQAEADLSSTTTLQRPQQA 7172
Db 1741 PCSSPFGSASQASSQVSSLRVSSQVGTGTEPGSLDAGWTQAEADLSSTTTLQRPQQA 1800
Qy 7173 TMRKFSLGRCGYAGVAGVTGTFAGDGAGMLGQGPWARIATAWVSQSEEESEAEARAE 7232
Db 1801 TMRKFSLGRCGYAGVAGVTGTFAGDGAGMLGQGPWARIATAWVSQSEEESEAEARAE 1860
Qy 7233 QSEEQEAREASPLQVVSARVPVEVGRAPTRSSPEPTPWEDIGQVSLVQIRDLSDADAA 7292
Db 1861 QSEEQEAREASPLQVVSARVPVEVGRAPTRSSPEPTPWEDIGQVSLVQIRDLSDADAA 1920
Qy 7293 DTISLSDISEVDPAYNLSDLYDKYLPPEFMIFRKVPKSAQPEPPSPMAEBELAEFFPEPT 7352
Db 1921 DTISLSDISEVDPAYNLSDLYDKYLPPEFMIFRKVPKSAQPEPPSPMAEBELAEFFPEPT 1980
Qy 7353 WPWPGELGPHAGLEITTESEEDVDALLAEAAVGRKKKSSPSRSLFHPFGRLPLDDEPAEL 7412
Db 1981 WPWPGELGPHAGLEITTESEEDVDALLAEAAVGRKKKSSPSRSLFHPFGRLPLDDEPAEL 2040
Qy 7413 GLRERVKASVEHISRILKGRPEGLEKEGPPRRKPGLASFRLSGLKSWDRAPFTFLRELSDE 7472
Db 2041 GLRERVKASVEHISRILKGRPEGLEKEGPPRRKPGLASFRLSGLKSWDRAPFTFLRELSDE 2100
Qy 7473 TVVLQGSVTLACQVSAQAAQATWSKQCAPLESSESRVLISATLKNFQLLTILVVAEDLG 7532
Db 2101 TVVLQGSVTLACQVSAQAAQATWSKQCAPLESSESRVLISATLKNFQLLTILVVAEDLG 2160
Qy 7533 VYTCVSNALGTVTTTGVLRKAERSSSPCPDIGEVYADGVLLVWKPVESYGPVTYIVQC 7592
Db 2161 VYTCVSNALGTVTTTGVLRKAERSSSPCPDIGEVYADGVLLVWKPVESYGPVTYIVQC 2220

Qy 7593 SLEGGSWTTLASDIFDCCYLTSKLSRGQTYTFRACYSKAGWGPYSSPSEQVLLGGPSHL 7652
Db 2221 SLEGGSWTTLASDIFDCCYLTSKLSRGQTYTFRACYSKAGWGPYSSPSEQVLLGGPSHL 2280
Qy 7653 ASEESQGRSAQPLPSTKTFQTOIQRGFRSVVRQWKEKASGRALAAKIIIPYHPKOKTA 7712
Db 2281 ASEESQGRSAQPLPSTKTFQTOIQRGFRSVVRQWKEKASGRALAAKIIIPYHPKOKTA 2340
Qy 7713 VLREYEALKGLRHPLHAQLHAAYLSRPHLVILILELCSGPELLPCLAEASYSSEVSKDYL 7772
Db 2341 VLREYEALKGLRHPLHAQLHAAYLSRPHLVILILELCSGPELLPCLAEASYSSEVSKDYL 2400
Qy 7773 WQMLSATQYLNHCHILHLDLRSENMIITEYNLLKVVDLGNQSLSOEKVLPDSDFKDYLE 7832
Db 2401 WQMLSATQYLNHCHILHLDLRSENMIITEYNLLKVVDLGNQSLSOEKVLPDSDFKDYLE 2460
Qy 7833 TMAPELLEGGQAVPQTDIWAIGVTFIMLSAEPVPSSEGARDLQRLKGLVRLSRCYAG 7892
Db 2461 TMAPELLEGGQAVPQTDIWAIGVTFIMLSAEPVPSSEGARDLQRLKGLVRLSRCYAG 2520
Qy 7893 LSGGAVAFPLSTLCAOPWGRPCASSCLOCQWLTTEGPACSRPAPVTPPTARLRFVVRRE 7952
Db 2521 LSGGAVAFPLSTLCAOPWGRPCASSCLOCQWLTTEGPACSRPAPVTPPTARLRFVVRRE 2580
Qy 7953 KRRALLYKRHNLAQVR 7968
Db 2581 KRRALLYKRHNLAQVR 2596
RESULT 5
ABP70084
ID ABP70084 standard; protein; 4691 AA.
XX AC ABP70084;
XX DT 27-JAN-2003 (first entry)
XX DE Human NOV13a.
KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
KW antiinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
KW neutropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
KW antifertility; cerebroprotective; gene therapy; NOVX; fertility;
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW immune disorder; hematopoietic disorder; cardiovascular disorder;
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis.
XX OS Homo sapiens.
XX PN WO200272771-A2.
XX PD 19-SEP-2002.
XX PF 08-MAR-2002; 2002WO-US007288.
XX PR 08-MAR-2001; 2001US-0274101P.
PR 08-MAR-2001; 2001US-0274194P.
PR 08-MAR-2001; 2001US-0274281P.
PR 08-MAR-2001; 2001US-0274322P.
PR 09-MAR-2001; 2001US-0274849P.
PR 12-MAR-2001; 2001US-0275235P.
PR 13-MAR-2001; 2001US-0275578P.
PR 13-MAR-2001; 2001US-0275579P.
PR 13-MAR-2001; 2001US-0275601P.
PR 14-MAR-2001; 2001US-0276000P.
PR 16-MAR-2001; 2001US-0276776P.
PR 19-MAR-2001; 2001US-0276994P.
PR 20-MAR-2001; 2001US-0277329P.
PR 20-MAR-2001; 2001US-0277321P.

PR 20-MAR-2001; 2001US-0277327P.
PR 20-MAR-2001; 2001US-0277338P.
PR 21-MAR-2001; 2001US-0277739P.
PR 22-MAR-2001; 2001US-0277783P.
PR 23-MAR-2001; 2001US-0278152P.
PR 26-MAR-2001; 2001US-0278899P.
PR 27-MAR-2001; 2001US-0278899P.
PR 27-MAR-2001; 2001US-0279036P.
PR 28-MAR-2001; 2001US-0279348P.
PR 30-MAR-2001; 2001US-0279999P.
PR 30-MAR-2001; 2001US-0280233P.
PR 02-APR-2001; 2001US-0280802P.
PR 02-APR-2001; 2001US-0280822P.
PR 02-APR-2001; 2001US-0280900P.
PR 04-APR-2001; 2001US-0281194P.
PR 13-APR-2001; 2001US-0283675P.
PR 13-APR-2001; 2001US-0287424P.
PR 02-MAY-2001; 2001US-0288066P.
PR 03-MAY-2001; 2001US-0288342P.
PR 15-MAY-2001; 2001US-0291190P.
PR 16-MAY-2001; 2001US-0291240P.
PR 30-MAY-2001; 2001US-0294485P.
PR 31-MAY-2001; 2001US-0294889P.
PR 31-MAY-2001; 2001US-0294899P.
PR 18-JUN-2001; 2001US-0299027P.
PR 19-JUN-2001; 2001US-0299303P.
PR 19-JUN-2001; 2001US-0299310P.
PR 10-JUL-2001; 2001US-0304354P.
PR 31-JUL-2001; 2001US-0309198P.
PR 16-AUG-2001; 2001US-0312903P.
PR 10-SEP-2001; 2001US-0318463P.
PR 12-SEP-2001; 2001US-0318770P.
PR 27-SEP-2001; 2001US-0325430P.
PR 14-NOV-2001; 2001US-0325681P.
PR 18-OCT-2001; 2001US-0330380P.
PR 31-OCT-2001; 2001US-0335301P.
PR 14-NOV-2001; 2001US-0332172P.
PR 14-NOV-2001; 2001US-0332271P.
PR 14-NOV-2001; 2001US-0332272P.
PR 14-NOV-2001; 2001US-0333184P.
PR 21-NOV-2001; 2001US-0333272P.
PR 03-DEC-2001; 2001US-0332094P.
PR 03-DEC-2001; 2001US-0337426P.
PR 04-DEC-2001; 2001US-0337185P.
PR 03-JAN-2002; 2002US-0345705P.
PR 08-MAR-2002; 2002US-00093463.
XX
FA (CURA-) CURAGEN CORP.
XX Rastelli L, Mezes PD, Smithson G, Guo X, Gerlach V, Casman SJ;
XX Boldog FL, Li L, Zerhusen BD, Tchernov VT, Gangolli EA, Vernet CM;
PI Pena CE, Bargas CE, Liu X, Spytek KA, Gorman L, Spaderna SK;
PI Voss EZ, Malyankar UM, Anderson DW, Patturajan M, Miller CE;
PI Taupier RJ, Padigar M, Shenoy SG, Kekuda R, Gusev VV, Pochart PF;
PI Zhong M;
XX WPI; 2002-732824/79.
XX N-PSDB; ABV99362.
XX
XX New NOVX polypeptides and polynucleotides, useful for preventing,
XX diagnosing or treating NOVX-associated disorders e.g. diabetes, cancer,
XX Alzheimer's disease, dyslipidemias, obesity, immune or hematopoietic
XX disorders, and asthma.
XX
XX Claim 1; Page 137-138; 619pp; English.
XX
XX The present invention relates to new isolated proteins (NOVX) and their
XX coding sequences (ABV99327-ABV99595 and ABP70049-ABP70149), where X is
XX any number from 1 to 48. The NOVX proteins and coding sequences are
XX useful in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, preferably a NOVX-associated disorder.
CC The NOVX coding sequences and proteins are useful for treating,
CC preventing or diagnosing diseases such as metabolic disorders, diabetes,
CC obesity, infectious disease, anorexia, cancer-associated cachexia,
CC cancer, neurodegenerative diseases, Alzheimer's disease, Parkinson's
CC disease, immune disorders, hematopoietic disorders, cardiovascular
CC disorders, fertility, bronchial asthma, AIDS, dyslipidemia, metabolic
CC disturbances associated with obesity, metabolic syndrome X or wasting
CC disorders associated with chronic diseases or various cancers. The NOVX
CC coding sequences and proteins may also be used as targets for the
CC identification of small molecules that modulate or inhibit e.g.
CC neurogenesis, cell differentiation, cell proliferation, hematopoiesis,
CC wound healing and angiogenesis, in gene therapy, in generation of
CC antibodies that bind immunospecifically to NOVX substances for use in
CC therapeutic or diagnostic methods
XX
XX SQ Sequence 4691 AA;
Query Match 25 5%; Score 10519.5; DB 5; Length 4691;
Best Local Similarity 51.1%; Pred. No. 0;
Matches 2367; Conservative 118; Mismatches 342; Indels 1807; Gaps 66;
QY 2595 MPLYNDSPHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTARSARLEVRVPVFLKALD 2654
DB 1 MPLYNDSPHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTARSARLEVRVPVFLKALD 60
QY 2655 DLSAERTGLALQCEVSDPEAHVVRKDGVLGSPDKYDFLHTAGTGLVHVHVPEDAG 2714
DB 61 DLSAERTGLALQCEVSDPEAHVVRKDGVLGSPDKYDFLHTAGTGLVHVHVPEDAG 120
QY 2715 LYTCHVGSEETRARVRVHDHLVGTIKRLKTMVELEGSCSECVLSHESADPAMTVGG 2774
DB 121 LYTCHVGSEETRARVRVHDHLVGTIKRLKTMVELEGSCSECVLSHESADPAMTVGG 180
QY 2775 KTVGSSRFQATROGRKYILVREAAPSDAGEVFSVRGLTSKASLIVRERPAATIKPLE 2834
DB 181 KTVGSSRFQATROGRKYILVREAAPSDAGEVFSVRGLTSKASLIVRERPAATIKPLE 240
QY 2835 DQVAPGEDVELCERAGTFVHVKDKRAIKRSQKYDVVCEGTMAMLVIRGASLKDAG 2894
DB 241 DQVAPGEDVELCERAGTFVHVKDKRAIKRSQKYDVVCEGTMAMLVIRGASLKDAG 300
QY 2895 EYTCEVASKSTASHVEEKANCFTETLNQVEEKGTAFTCTKTEHPAATVTRKGLLE 2954
DB 301 EYTCEVASKSTASHVEEKANCFTETLNQVEEKGTAFTCTKTEHPAATVTRKGLLE 360
QY 2955 LRASGKHQPSQEGTLRLTISALEKADSDTYTCDIGQAQSAQLLVQRRVHIIDLEDV 3014
DB 361 LRASGKHQPSQEGTLRLTISALEKADSDTYTCDIGQAQSAQLLVQRRVHIIDLEDV 420
QY 3015 DVQEGSSATPRCISPANYEPVHWFLLDKTPLHANELNEIDAOPGGYHVLTILRLQALKDSG 3074
DB 421 DVQEGSSATPRCISPANYEPVHWFLLDKTPLHANELNEIDAOPGGYHVLTILRLQALKDSG 480
QY 3075 TIYFAGDQASAAALRTERKPSVRESRLTDATITEGEDTLVCTSTCDIPMCWTGDKT 3134
DB 481 TIYFAGDQASAAALRTERKPSVRESRLTDATITEGEDTLVCTSTCDIPMCWTGDKT 540
QY 3135 LRGSARCOLSHEGHRQALLITGATLQDSGRYKCBAGGACSSSIVRVHARVPRFQALKDL 3194
DB 541 LRGSARCOLSHEGHRQALLITGATLQDSGRYKCBAGGACSSSIVRVHARVPRFQALKDL 600
QY 3195 EVLEGGATLRCLVSSVAAPVKWCYGNVLRPGDKYSLRQEGAMLELVVRLRPQDSGRY 3254
DB 601 EVLEGGATLRCLVSSVAAPVKWCYGNVLRPGDKYSLRQEGAMLELVVRLRPQDSGRY 660
QY 3255 SCSFGDQTTATLTVTALPAQFICKLRNKEGATATLRCELSKTAPVEVRKGSSETLRD 3314
DB 661 SCSFGDQTTATLTVTALPAQFICKLRNKEGATATLRCELSKTAPVEVRKGSSETLRD 720
QY 3315 GDRYCLQDQGMCELQIRGLAWDAAEYSCVCGERTSASLTIRPMPAHPIGLRLRHQSSI 3374
DB 721 GDRYCLQDQGMCELQIRGLAWDAAEYSCVCGERTSASLTIRPMPAHPIGLRLRHQSSI 780

Db 2877 QDNVVARLTFRPALPCDSIGSIYCEAAGTRVVALLOVQAKNTVVRRGLENVLEGEALFE 2936
|||
Qy 4361 COLSQPEVAHATWLLDDBPVRTSENAEVVFFENGLRHLLLLKNLRPODSQRYTFILAGDMV 4420
|||
Db 2937 COLSQPEVAHATWLLDDBPVRTSENAEVVFFENGLRHLLLLKNLRPODSQRYTFILAGDMV 2996
|||
Qy 4421 TSAFLTVR-----GHRLEILEPLKNAAVRAGAQARFTCTLSAIVPVGEASWINGAA 4472
|||
Db 2937 TSAFLTVRGCAVLVQGRLEILEPLKNAAVRAGAQARFTCTLSAIVPVGEASWINGAA 3056
|||
Qy 4473 VQPDSDMTVTADSGHQAALLRSACOPHRAGEVTFACRDVAASARLTVLGLDPDPDAEVV 4532
|||
Db 3057 VQPDSDMTVTADSGHQAALLRSACOPHRAGEVTFACRDVAASARLTVLGLDPDPDAEVV 3116
|||
Qy 4533 AHSSHTVTLSWAAPMSDGGGLCGYRVEVEKATQWRMLCHELVPGPECVVDGLAPGETY 4592
|||
Db 3117 ARSSHTVTLSWAAPMSDGGGLCGYRVEVEKATQWRMLCHELVPGPECVVDGLAPGETY 3176
|||
Qy 4593 RFRVAAGVGVGAGEPVHLPTQVRLAEPKVPVPPQPSAPESRQVAAGEDVLSLEVVAEAG 4652
|||
Db 3177 RFRVAAGVGVGAGEPVHLPTQVRLAEPKVPVPPQPSAPESRQVAAGEDVLSLEVVAEAG 3235
|||
Qy 4653 EVIWHKMERIQPGRFVWSQGRQMLVIKGFATDQGEVHCGLAQGSICPAAATFOVA 4712
|||
Db 3236 EVIWHKMERIQPGRFVWSQGRQMLVIKGFATDQGEVHCGLAQGSICPAAATFOVA 3295
|||
Qy 4713 LSPASVDEAPQPSLPPEAAQEGDLHLLWEALARKRMSREPTLDSISELPEEDGRSQRLP 4772
|||
Db 3296 LSPASVDEAPQPSLPPEAAQEGDLHLLWEALARKRMSREPTLDSISELPEEDGRSQRLP 3355
|||
Qy 4773 QEAERVADELSEGYSTADELARTGADLSHTSSDDSRAGTSPSLVLYLKAGRPQTSPLA 4832
|||
Db 3356 QEAERVADELSEGYSTADELARTGADLSHTSSDDSRAGTSPSLVLYLKAGRPQTSPLA 3415
|||
Qy 4833 SKVGAP-----RLLTKVAEMA-----RLLTKVAEMA----- 3568
|||
Db 3416 SKVSPNLAKEFRFTPRAGRSLGLFVGADPAFPGSERSARTRCAAPPRESUKREP- 3474
|||
Qy 4853 AAVRPLGLDSTKDLG-----DPSMD-----KAAVKIQ-----AAFK 4884
|||
Db 3475 ASCLP--GAMEAVELARKLQBEATCSICLDYFTDPVMTTCGHNFCEACIQLSWEKARGK 3532
|||
Qy 4895 GYKVRK-----EMKQOGPMFSGHTGTEAQVGDALRLCEVCVASKADVRAWLKDG 4935
|||
Db 3533 GRKRGKGFPCPECREMSPQNLFP-----RLLTKVAEMA----- 3568
|||
Qy 4936 VELTDGRHHHIDQDGTCSLLIAGLDRADAGCYTCOVSNK-----FGQVTHSACVVSQS 4991
|||
Db 3569 -----QGH-----PGLQKOD-----LCQBHHBPLKLFCKQKQSPICVVCRE 3604
|||
Qy 4992 ESEAESSGGELDAPFRBAARLHRLPRTKSPAESVDEELFLSADEG----- 5038
|||
Db 3605 SRE-----HRLHVL-----PAEAVOGYKLUKEEDMEYLRQITRTGN 3643
|||
Qy 5039 --PAPPEPAMQ--TYRDEHPICIRFAL-----TEARQAVTRFQMF 5079
|||
Db 3644 LQAREQSLAEQVQKVRERRIVLEFEKMNLYLVEERQRLIQALETTEETASRLRSV 3703
|||
Qy 5080 ATL--GIGVEIKLVE-----QGRPR-----VEMCISK-- 5104
|||
Db 3704 ACLDRQSHLELLQLLBERSTQGPLQMLQDMKEPLSRAALLVLIHGNLVEFPVSLP 3763
|||
Qy 5105 -----ETPAPVVPPEPLSLTSDAAP-----VFLTE 5131
|||
Db 3764 SPLYLIATKAHTQLGPGTTPDPECTPLPISPP--PRES--TEDVVPDATSAYPVLLIYE 3820
|||
Qy 5132 LQNOBQVQGYP-----VSPDCVVTPGQPMPSVRWFKDGKLLBEDDHYMINEDQOG 5180
|||
Db 3821 SRQRRLGSSPEGSGFCSDRFPVAVPCAV--GQTA-----FSSGR-----HYWEVGMNIT 3868
|||
Qy 5191 GHQLIITAVPADMG-----VVRCLAENSMGVSSTKAELRVLTSTDYTDAADESSSYF 5236
|||

Db 3869 GDALWALGVCRDNVSRKDRVPKC--PENGFWV-----VOLSK-----GTYKLSLF 3911
|||
Qy 5237 SAQGYLSRE--OEGTESTTDEGQLPOVVEELRLDQVAGTRLAKFOLKVKGYPAPRLY 5293
|||
Db 3912 SALTPVLMLEPPSHMGIFLDFEAG-----EVSFVSVDGSHLHTYQAT--PPGPLQP 3962
|||
Qy 5294 WFKGQPLTASAHIRMTCKKILHTLEIISVTREDSGQYAAVYISNMGAAYSARLLVRGP 5353
|||
Db 3963 FFCLGAP-----KSCQMVISTVMAGV--KDLATRTGAVVTPALGA----- 4001
|||
Qy 5354 DEPEEKPADVHEQLVPPMLERFTPKVKKKGSSITFSVKVEGRVPVTVHMLREBAERG 5413
|||
Db 4002 -----YAPSATEYQS-----PAP--WSPRAPE-- 4021
|||
Qy 5414 LWIGDPTPGYTVASSAQOHSVLVLDVGRHQHGYTYTCIASNAAGQALCSLHVSGLPKVE 5473
|||
Db 4022 ----PEHPG--VPSLAPRS-----RACAAAPGYPGSPRAA 4051
|||
Qy 5474 EQEKVEKALISTFLOCTTOAISAOGLETASADILGGQRKEEPLAKEALHLSLAEVGT 5533
|||
Db 4052 EAARRRPADSTAFILP--SVRAMAA-----PDLSTNLQEBEATCA----- 4087
|||
Qy 5534 EFLOQLVTSQITVMVSAKITQAKLOVPVGGSDSDSKTSPAS--PRHGRSRPSSSI----- 5585
|||
Db 4088 ICLDYFTDPVMTDCGHNFCEIRCWGQEPARTAPSAASCPRGTCGTARTLLRWPRW 4147
|||
Qy 5586 QESSESESDGARGEIEDIYVVTADYLPGLAE--QDAITL-----REGQYVEVLDAAHPL 5638
|||
Db 4148 RGACTRRRRRR-----VPAHREPLAAFCGDELRLCAACERSGEH----- 4189
|||
Qy 5639 RWL-----VRTKPKSSPSROG-----WVSPAYLDRLK--LSPE----- 5671
|||
Db 4190 --WAHVRGCRTRPKTSRPLEAGTMAANETLSGAKLEKSLHRLKQMODALLFQAQDET 4248
|||
Qy 5672 ---NGAAEAPFPCEAVSEDEYKARLSVIOELLSSQAFVEELQFLOSHHL---QHLE 5724
|||
Db 4249 CVLQHAEDGGQRQNVLEFE---RLRLLEAGGTAATAAAGEBELKQSAHLAELAELE 4305
|||
Qy 5725 RCPHVPIAVAGOKA 5738
|||
Db 4306 R--PLPAACAGAAA 4317
|||
RESULT 6
ABP70085
ID ABP70085 standard; protein; 4675 AA.
XX AC ABP70085;
XX DT 27-JAN-2003 (first entry)
XX DE Human NOV13b.
XX KW Human; anti-HIV; cytostatic; antidiabetic; antiasthmatic; cachexia; AIDS;
KW antinflammatory; cardiant; haemostatic; neuroprotective; anorectic;
KW nontropic; immunosuppressive; osteopathic; antiparkinsonian; cancer;
KW antifertility; cerebroprotective; gene therapy; NOV; fertility;
KW metabolic disorder; diabetes; obesity; infectious disease; anorexia;
KW neurodegenerative disease; Alzheimer's disease; Parkinson's disease;
KW immune disorder; haematopoietic disorder; cardiovascular disorder;
KW bronchial asthma; dyslipidemia; metabolic disturbance; neurogenesis;
KW metabolic syndrome X; wasting disorder; cell differentiation;
KW cell proliferation; haematopoiesis; wound healing; angiogenesis.
XX OS Homo sapiens.
XX PN WO200272771-A2.
XX PD 19-SEP-2002.
XX PF 08-MAR-2002; 2002WO-US007288.
XX PR 08-MAR-2001; 2001US-0274101P.

Db 541 LRGSARCOLSHBGHRAQLLITGATLQDSGRYKCEAGGACSSSIVRVHARPVRFQEAIXDL 600
Qy 3195 EVLEGGAATLRCLVSSVAAPVWCVGNVLRPGDKYSLRQBGAMLELVVRNLRPQDSGRY 3254
Db 601 EVLEGGAATLRCLVSSVAAPVWCVGNVLRPGDKYSLRQBGAMLELVVRNLRPQDSGRY 660
Qy 3255 SCSFGDQTTSATLTVTALPAQPIGKLRNKEATEGATATLRCELSKTAPVWVRKGSSETLRD 3314
Db 661 SCSFGDQTTSATLTVTALPAQPIGKLRNKEATEGATATLRCELSKTAPVWVRKGSSETLRD 720
Qy 3315 GDYRCLRDQGMACELOIRGLAMVDAEYSVCVGEERTSASLTIRPMPAHFIRGLRHQESI 3374
Db 721 GDYRCLRDQGMACELOIRGLAMVDAEYSVCVGEERTSASLTIRPMPAHFIRGLRHQESI 780
Qy 3375 EGATATLRCELSKAAPVWVRKGRRESLRDGRHSRLRQDGAVALCELOICGLAVADAGEYSVCV 3434
Db 781 EGATATLRCELSKAAPVWVRKGRRESLRDGRHSRLRQDGAVALCELOICGLAVADAGEYSVCV 840
Qy 3435 GEERTSATLTVKALPAKFTTEGLRNEAEVAGATAMLCESKVAPVWVRKGPENLRDGRY 3494
Db 841 GEERTSATLTVKALPAKFTTEGLRNEAEVAGATAMLCESKVAPVWVRKGPENLRDGRY 900
Qy 3495 ILRQSGTRCELOICGLAMADAGEYLCVCGQERTSATLTIRALPARFIEDVKNQAEAREGAT 3554
Db 901 ILRQSGTRCELOICGLAMADAGEYLCVCGQERTSATLTIRALPARFIEDVKNQAEAREGAT 960
Qy 3555 AVLOCELNSAAPVWVRKGSSETLRDGRYSLRQDGTKELOIRGLAMADTGEYSVCVCGQER 3614
Db 961 AVLOCELNSAAPVWVRKGSSETLRDGRYSLRQDGTKELOIRGLAMADTGEYSVCVCGQER 1020
Qy 3615 TSAMLTVRALPIKFTTEGLRNEATEGATAVLRCELSKAPVWVRKGHETLRDGRHSRLRQ 3674
Db 1021 TSAMLTVRALPIKFTTEGLRNEATEGATAVLRCELSKAPVWVRKGHETLRDGRHSRLRQ 1080
Qy 3675 DGARCELOIRGLVADAGEYLCMCKERTSAMLTVRAMPSPKFI EGLRNEEATEGDTATLW 3734
Db 1081 DGARCELOIRGLVADAGEYLCMCKERTSAMLTVRAMPSPKFI EGLRNEEATEGDTATLW 1140
Qy 3735 CELSKAAPVWVRKGHETLRDGRHSRLRQDGRSCELOIRGLAVVDAGEYSVCVCGQERTSAT 3794
Db 1141 CELSKAAPVWVRKGHETLRDGRHSRLRQDGRSCELOIRGLAVVDAGEYSVCVCGQERTSAT 1200
Qy 3795 LTVRALPARFIEDVKNQAEAREGATAVLOCELSKAAPVWVRKGSSETLRDGRYSLRQDGR 3854
Db 1201 LTVRALPARFIEDVKNQAEAREGATAVLOCELSKAAPVWVRKGSSETLRDGRYSLRQDGR 1260
Qy 3855 CELOIHGLSVADTGEYSVCVCGQERTSATLTVR----- 3886
Db 1261 CELOIHGLSVADTGEYSVCVCGQERTSATLTVRALPARFTQDLKTKEASEGATATLQCELS 1320
Qy 3887 ----- 3886
Db 1321 KVAPVWVRKGPETLRDGRYSKQDGTCELOIHDLVSADAGEYSVCVCGQERTSATLTVR 1380
Qy 3887 ----- 3886
Db 1381 ALPARFTEGLRNEEAMEGATATLQCELSKAAPVWVRKGLEALRDGDKYSLRQDGAVALCELO 1440
Qy 3887 ----- 3886
Db 1441 IHGLAMADGVYSCVCGQERTSATLTVRALPARFIEDVKNQAEAREGATVTVLQCKLRKAAP 1500
Qy 3887 ----- 3886
Db 1501 VEWKGPNTLKQDGRYSKQDGTSCLOIRGLVIADAGEYSVCVCGQERTSATLTVRALPA 1560
Qy 3887 ----- 3886
Db 1561 RFIEDVRNHEATEGATAVLOCELSKAAPVWVRKGSSETLRDGRYSLRQDGTCELOIRGL 1620
Qy 3887 ----- 3886
Db 1621 AVEDTGEYLCVCGQERTSATLTVRALPARFIDNMTNQEAREGATATLHCELSKVAPVWVR 1680

Qy 3887 ----- 3886
Db 1681 KGPETLRDGRHSRLRQDGRSCELOIRGLAVVDAGEYSVCVCGQERTSATLTVRALPARFIE 1740
Qy 3887 ----- 3886
Db 1741 DVKNQAEAREGATAVLOCELSKAAPVWVRKGSSETLRDGRYSLRQDGTCELOIHGLSVAD 1800
Qy 3887 ----- 3886
Db 1801 TGEYSVCVCGQERTSATLTVRALPARFTQDLKTKEASEGATATLQCELSKVAPVWVRKGP 1860
Qy 3887 ----- 3886
Db 1861 TLRDGRYSKQDGTCELOIHDLVSADAGEYSVCVCGQERTSATLTVRDCHTLHVMPHY 1920
Qy 3887 ----- 3886
Db 1921 FQLPGLLKEPETLIYIQIPSPVILFTEGLRNEEAMEGATATLQCELSKAAPVWVRKGL 1980
Qy 3887 ----- 3886
Db 1981 ALRDGDKYSLRQDGAVALCELOIHGLAMADGVYSSLPARFIEDMRNOKATEGATVTVLQCKL 2040
Qy 3887 ----- 3886
Db 2041 RKAAPVWVRKGPNTLKQDGRYSKQDGTSCLOIRGLVIADAGEYSVCVCGQERTSATLTV 2100
Qy 3887 ----- 3886
Db 2101 RALPARFIEDVRNHEATEGATAVLOCELSKAAPVWVRKGSSETLRDGRYSLRQDGTCE 2160
Qy 3887 ----- 3886
Db 2161 QIRGLAVEDTGEYLCVCGQERTSATLTVRALPARFIDNMTNQEAREGATATLHCELSKVA 2220
Qy 3887 ----- 3886
Db 2221 PVWVRKGPETLRDGRHSRLRQDGTCELOIRGLSVADAGEYSVCVCGQERTSATLTVRALP 2280
Qy 3887 ----- 3886
Db 2281 AKFTKGLRNEEATEGATAMLOCELSKVAPVWVRKGPETLRDGRYSLRQDGTCELOIHG 2340
Qy 3887 ----- 3886
Db 2341 LSVADTGEYSVCVCGQERTSATLTVKAPQPVFPFREPQLQAESEGSTATLQCELSSEPTATV 2400
Qy 3922 WSKGGLQLOANGRRREPRLQGGCTABELVLQDLQREDTGEYTCCTCGSQATSATLTVTAAPVR 3981
Db 2401 WSKGGLQLOANGRRREPRLQGGCTABELVLQDLQREDTGEYTCCTCGSQATSATLTVTAAPVR 2460
Qy 3982 LRELQHOEVDSEGGTAHLCCELSRAGASVWVRKGSLOLPPCAKYOMVODGAAELLRVGE 4041
Db 2461 LRELQHOEVDSEGGTAHLCCELSRAGASVWVRKGSLOLPPCAKYOMVODGAAELLRVGE 2520
Qy 4042 QEDAGDYTCDTGHTQSMASLSVRVPRPKFTRLSQLEOETGDIARLCCOLSDABESGAVVQ 4101
Db 2521 QEDAGDYTCDTGHTQSMASLSVRGGR-----GAA-- 2549
Qy 4102 WLKSGVELHAGPKYKEMRSQGATRELLIHQLBAKOTGEYACVYTGOKTAASLRVTEPEVTI 4161
Db 2550 -----CGPVRDAAQCATRELLIHQLEAKDTGEYACVYTGOKTAASLRVTEPEVTI 2600
Qy 4162 VRGLVDAEVTDAEDVEFSCEVSRAGATGVQCLQGLPLQSNVEVTEVAVRDRGRIHTRLKG 4221
Db 2601 VRGLVDAEVTDAEDVEFSCEVSRAGATGVQCLQGLPLQSNVEVTEVAVRDRGRIHTRLKG 2660
Qy 4222 VTPEDAGTVSFHLGNHASSAQLTVRAPEVTITILEPLQDQVQLSEGQDASFCRLSRASGOEA 4281
Db 2661 VTPEDAGTVSFHLGNHASSAQLTVRAPEVTITILEPLQDQVQL----- 2701

QY 4282 RWAAGVPLQANEMNDITVEQGTLLHLLTLHKVTLTDAGTVSPHVGTCSSEAQLKVT----- 4337
DB |||||
QY 2702 -----GVPLQANEMNDITVEQGTLLHLLTLHKVTLTDAGTVSPHVGTCSSEAQLKVT EAVP 2756
DB |||||
QY 4338 ----- 4337
DB 2757 CLVRLQNVDFPAGEVATFSCBDGPQSAIAVRDGIHFHLSLGLGVADSGTVIFRAGPLV 2816
QY 4338 ----- 4337
DB 2817 STAKLLIKDPVVEVVSAMQDLAVEEGGSALLCQYSRPVQATWKMDEREVHTDGRHVIE 2876
QY 4338 -----AKNTVVRGLNVEALEGGEALFE 4360
DB 2877 QDNVRLTFRPAIPCDSGIYSCEAAGTRVALLVQVQAKNTVVRGLNVEALEGGEALFE 2936
QY 4361 COLSQEVAHAHTWLLDDPVRTSENAEVVFFENGRLHLLLLKNLRPQDSQCRVTFLAGDMV 4420
DB |||||
QY 2937 COLSQEVAHAHTWLLDDPVRTSENAEVVFFENGRLHLLLLKNLRPQDSQCRVTFLAGDMV 2996
QY 4421 TSAFLTVR-----GWRLEILEPLKNAAVRAGAQAQRTCTLSEAVPVGEASWYINGAA 4472
DB |||||
QY 2997 TSAFLTVRGDCAVLVQGRWLEILEPLKNAAVRAGAQAQRTCTLSEAVPVGEASWYINGAA 3056
QY 4473 VQPDSDMTVTADGSHQALLLSAQPHHAGEVTFACRDVAVASARLTVLGLPDPPEDAEVV 4532
DB |||||
QY 3057 VQPDSDMTVTADGSHQALLLSAQPHHAGEVTFACRDVAVASARLTVLGLPDPPEDAEVV 3116
QY 4533 AHSSTHTVLSWAAPMSDGGGLCGYRVEVKEGATQWRLCHELVPGPCVDGLAPGTY 4592
DB |||||
QY 3117 AHSSTHTVLSWAAPMSDGGGLCGYRVEVKEGATQWRLCHELVPGPCVDGLAPGTY 3176
QY 4593 RFRVAAGPVGAGEPVHLPQTVRLAEPKPVPPQPSAPESROVAAGEDVSLSELEVVAEAG 4652
DB |||||
QY 3177 RFRVAAGPVGAGEPVHLPQTVRL-EPKPVPPQPSAPESROVAAGEDVSLSELEVVAEAG 3235
QY 4653 EVIWHKGMRIQPGGRFVVSQGRQOVLVIKGFATDQGEYHCGLAQGSICPAAATFOVA 4712
DB |||||
QY 3236 EVIWHKGMRIQPGGRFVVSQGRQOVLVIKGFATDQGEYHCGLAQGSICPAAATFOVA 3295
QY 4713 LSPASVDEAPQPSLPEAAQEGDLHLLWEALARKRMSREPTLDSISLPEPDEGRSRLP 4772
DB |||||
QY 3296 LSPASVDEAPQPSLPEAAQEGDLHLLWEALARKRMSREPTLDSISLPEPDEGRSRLP 3355
QY 4773 QBAEVAAPDLSEGYSTADELARTGADLSHTSSDDSRAGTSPSLVYLKKAAGRPGTSPLA 4832
DB |||||
QY 3356 QBAEVAAPDLSEGYSTADELARTGADLSHTSSDDSRAGTSPSLVYLKKAAGRPGTSPLA 3415
QY 4833 SKVGAP-----AAPSVKPQQQOEPL 4852
DB |||||
QY 3416 SKVSPNLACKERFPTPRAGRSLLGVGADPAFPGERSARCTRRCAAPPPRESLKRBP- 3474
QY 4853 AAVRPLGLSLTKDGL-----DPSMD-----KAAVKIQ-----AAPK 4884
DB |||||
QY 3475 ASCLP--GAMEAVEALARKLOBEATCSICLDVFTDPVMTTCGHNFRCRACIQLSWEKARGK 3532
QY 4885 GYVKRK-----EMKQKQGPMSFHTFGDTEAQVGDALRLCECVASKADVRLWKDG 4935
DB |||||
QY 3533 GRKKEGSPFPCPECKEMSPQNLFPN-----RLTKVAEMA----- 3568
QY 4936 VELTDGRHHHIDQLGDGTCSSLIAGLDRADAGCYTCQVSNK-----FGQVTHSACVVVSGS 4991
DB |||||
QY 3569 -----QGH-----PGLQKQD-----LCQEHHEPLKLFCKQKQSPICVVCRE 3604
QY 4992 ESEABSSGGELDDAFRAARLHFLTKSPAESVDEELFSLADEG----- 5038
DB |||||
QY 3605 SRE-----HRLHVL-----PABEAVOGYKLKEEDMEYLRBQITRTGN 3643
QY 5039 ---PAPEEPADWQ--TYREDEHFICIRFEAL-----TEARQAVTRFQEMF 5079
DB |||||
QY 3644 LOAREEQSLAEWQGVKERRERIVLEFEKMNLYLVEERQRLLOALETEEBETASRLRESV 3703
QY 5080 ATL-----GIGVEIKLVE-----QGPR-----VEMCISK-- 5104

DB 3704 ACLDRQGHSLLELLQLLEERSTQGPLQMLQDMKEPLSRAALLVVLHGMNLVEFPVSLP 3763
QY 5105 -----ETPAPVVPPEPLPSLLTSDAAP-----VFUTE 5131
DB 3764 SPLYLIAATKAHTQLGPGTTFDPCECTPLPISPP-PRPS--TEDVVPDATSAYPYLLLYE 3820
QY 5132 LQNEVQDGYP-----VSFDCVVTGQPMPSVRWFKDGKLLLEDDHY 5172
DB 3821 SRQRYLGSSPEGSGFCFKRFAVYPCAV-GQTA-----FSSGR-----HY 3860
RESULT 7
ID ABP58227 standard; protein; 2328 AA.
XX ABP58227;
AC ABP58227;
XX 31-MAR-2003 (first entry)
XX Human cell adhesion and extracellular matrix protein 4.
XX Cell adhesion and extracellular matrix protein 4; CADECM-4; human;
KW anti-HIV; virucide; antifungal; antiinflammatory; antianemic;
KW antiparkinsonian; nootropic; anticonvulsant; antifertility;
KW antiarteriosclerotic; antiasthmatic; immunosuppressive; antithyroid;
KW cytotatic; hepatotropic; dermatological; antidiabetic; nephrotropic;
KW angiot; thyromimetic; neuroprotective; osteopathic; antiarthritic;
KW antiparasitic; antihelminthic; antiparasitic; uropathic; ophthalmological;
KW antirheumatic; haemostatic; antibacterial; protozoacide; fungicide;
KW gynaecological; neurexin; gene therapy.
XX Homo sapiens.
XX WO200288322-A2.
XX 07-NOV-2002.
XX 01-MAY-2002; 2002WO-US013874.
XX 02-MAY-2001; 2001US-0288290P.
PR 21-MAY-2001; 2001US-0292468P.
PR 15-JUN-2001; 2001US-0298616P.
PR 28-JUN-2001; 2001US-0301672P.
PR 04-JAN-2002; 2002US-0345008P.
XX (INCY-) INCYTE GENOMICS INC.
XX Yue H, Lee EA, Duggan BM, Thangavelu K, Honchell CD, Ding L;
PI Hillman JL, Baughn MR, Kallick DA, Lee S, Warren BA, Xu Y, Tran UK;
PI Lal PG, Thornton M, Hafalia AJA, Yao MG, Nguyen DB, Gandhi AR;
PI Khan FA, Walhia NK, Griffin JA, Chinn AM, Elliott VS, Ramkumar J;
PI Arvizu CS, Forsythe LJ;
XX WPI; 2003-167112/16.
DR N-PSDB; ABZ24581.
XX New human cell adhesion and extracellular matrix proteins, useful for
PT diagnosing, treating or preventing autoimmune or inflammatory disorder
PT (e.g. AIDS, allergy or anemia), multiple sclerosis, osteoarthritis,
PT cancer or hepatitis.
XX Claim 1; Page 144-150; 178pp; English.
XX The present sequence is the protein sequence of Incyte polypeptide
CC 726129Cbl denoted human cell adhesion and extracellular matrix protein 4
CC (CADECM-4). The protein is encoded by a clone isolated from a male muscle
CC cDNA library. Homology searches indicate it to be a titin muscle protein.
CC The invention provides CADECM-1 to -11 polypeptides (see ABP58224-34) and
CC polynucleotides (see ABZ24578-88), expression vectors, host cells,
CC antibodies, agonists and antagonists. These are useful for diagnosing,
CC treating or preventing disorders associated with aberrant expression of
CC CADECM, particularly cell proliferative disorders (e.g. arteriosclerosis,

CC atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal
CC haemoglobinuria, polycythaemia vera, psoriasis, primary
CC thrombocytopenia or cancer), developmental disorders (e.g. renal
CC tubular acidosis, anaemia or mental retardation), neurological disorders
CC (e.g. Alzheimer's disease, Parkinson's disease or epilepsy), reproductive
CC disorders (e.g. infertility or a disruption in the menstrual cycle), or
CC autoimmune/inflammatory disorders (e.g. AIDS, allergy, asthma, autoimmune
CC thyroiditis, contact dermatitis, Crohn's disease, diabetes mellitus,
CC glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease,
CC Hashimoto's thyroiditis, irritable bowel syndrome, multiple sclerosis,
CC osteoarthritis, osteoporosis, pancreatitis, Reiter's syndrome, rheumatoid
CC arthritis, Sjogren's syndrome, uveitis), or vital, bacterial, fungal,
XX parasitic, protozoal or helminthic infections
SQ Sequence 2328 AA;

Query Match 23.8%; Score 9834.5; DB 6; Length 2328;
Best Local Similarity 84.6%; Pred. No. 0;
Matches 1941; Conservative 76; Mismatches 213; Indels 65; Gaps 8;

QY 2096 MEVQLSHADVDSWTRDGLRFGQPTCHLAVRGPMTLTLGLRPEDSGLMVFKAEVHT 2155
DB 1 MEVQLSHADVEGSWTRDGLRFGQPTCHLAVRGPMTLTLGLRPEDSGLMVFKAEVHT 60

QY 2156 SARLVVTELPSFSRPLQDVVTEKEKVTLECELSRPNVDVRLKDGVELRAGKTMATAA 2215
DB 61 SARLVVTELPSFSRPLQDVVTEKEKVTLECELSRPNVDVRLKDGVELRAGKTMATAA 120

QY 2216 QGACRSLLTYRCEPADQGVVYCDADHAQSSASVKVQ----- 2251
DB 121 QGACRSLLTYRCEPADQGVVYCDADHAQSSASVKVQGRNIQIVRPLEDEVMEKDGATFS 180

QY 2252 -----GRYTYLIVRRVLAEDAGBIQFVAENAESRA 2281
DB 181 CEVSHDEVPQWFEGSKLRPTDNVRIQEGRTYTLIVRRVLAEDAGBIQFVAENAESRA 240

QY 2282 QLRVKELPVTLRPLRDKTAMEKHKGVLCEQVSRASQVRVFKGSOELQPGPKYELVSDG 2341
DB 241 QLRVKELPVTLRPLRDKTAMEKHKGVLCEQVSRASQVRVFKGSOELQPGPKYELVSDG 300

QY 2342 LYRKLIISDVHAEDEDTYTCADGVKTSQAFFVEEQSITIVRGLQDVTMEPAPAFCE 2401
DB 301 LYRKLIISDVHAEDEDTYTCADGVKTSQAFFVEEQSITIVRGLQDVTMEPAPAFCE 360

QY 2402 TSIPSVRPPKLLGKTVLQAGNVGLEQEGTVHRLMLRRTCTMTGPHVFTVGKSRSSAR 2461
DB 361 TSIPSVRPPKLLGKTVLQAGNVGLEQEGTVHRLMLRRTCTMTGPHVFTVGKSRSSAR 420

QY 2462 LVVSDIPVLTPLRPLEPKTGRELQSVLSCDPRPAPKAVQWYKDDTPLSPSEKFKMSLEGQ 2521
DB 421 LVVSDIPVLTPLRPLEPKTGRELQSVLSCDPRPAPKAVQWYKDDTPLSPSEKFKMSLEGQ 480

QY 2522 MAELRILMPADAGVYRCQAGSAHSSTVTEAREVTVTGPDAAEATEEGWASFSCEL 2581
DB 481 MAELRILMPADAGVYRCQAGSAHSSTVTEAREVTVTGPDAAEATEEGWASFSCEL 540

QY 2582 SHEDDEVWSLNGMPLYNDSFHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTSARLE 2641
DB 541 SHEDDEVWSLNGMPLYNDSFHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTSARLE 600

QY 2642 VRVKPVVFLKALDDLISABERGTLALQCEVSDPEAHVWRKQVQLGPKDKYDFLHTAGTR 2701
DB 601 VRVKPVVFLKALDDLISABERGTLALQCEVSDPEAHVWRKQVQLGPKDKYDFLHTAGTR 660

QY 2702 GLVHVDVSPDAGLVTCVHGSSETRARVRVHDLHVGITKRLKTMVELEGESCSFPCVLSH 2761
DB 661 GLVHVDVSPDAGLVTCVHGSSETRARVRVHDLHVGITKRLKTMVELEGESCSFPCVLSH 720

QY 2762 ESADPAMWTVGKTVGSSSRFQATROGRKYTLVVREAAPS DAGBWFVSVRGLTSKASLI 2821
DB 721 ESADPAMWTVGKTVGSSSRFQATROGRKYTLVVREAAPS DAGBWFVSVRGLTSKASLI 780

QY 2822 VRERPAALIKPLEDQWVAPGEDVELURCSELRSAGTPVHMLKORKATRKQKYDVVCEGMA 2881

DB 781 VRERPAALIKPLEDQWVAPGEDVELURCSELRSAGTPVHMLKORKATRKQKYDVVCEGMA 840
QY 2882 MLVIRGASLKDAGEYTCVEASKSTASLHVBEKANKCFTEELTNLQVEBKGTAVFTCKTEH 2941
DB 841 MLVIRGASLKDAGEYTCVEASKSTASLHVBEKANKCFTEELTNLQVEBKGTAVFTCKTEH 900

QY 2942 PAATVTRWKGLELLELRASCKHQPQSBGLTLRLTISALEKADSDTYTCDTGOASRAQLLVQ 3001
DB 901 PAATVTRWKGLELLELRASCKHQPQSBGLTLRLTISALEKADSDTYTCDTGOASRAQLLVQ 960

QY 3002 GRRVHHIIEDLDDVDVQEGSSATFCRISPNANYPVHWFELDKTPLHANELNEIDAOPGGVH 3061
DB 961 GRRVHHIIEDLDDVDVQEGSSATFCRISPNANYPVHWFELDKTPLHANELNEIDAOPGGVH 1020

QY 3062 VLTLRQLALKDSGTIYFEAGDQASALRLVTEKPSVFSRELTDTATITEGEDTLVLCETST 3121
DB 1021 VLTLRQLALKDSGTIYFEAGDQASALRLVTEKPSVFSRELTDTATITEGEDTLVLCETST 1080

QY 3122 CDIPLMCWTGDKTILRGSAARCOLSHEGHRAQLITGATLQDSGRYKCEAGGACSSSIVRVH 3181
DB 1081 CDIPLMCWTGDKTILRGSAARCOLSHEGHRAQLITGATLQDSGRYKCEAGGACSSSIVRVH 1140

QY 3182 ARPVRFOALKDXLEVLGGAATLRCVLSSVAAPVKWCYGNVLRPGDKYSLRQEGAMLEL 3241
DB 1141 ARPVRFOALKDXLEVLGGAATLRCVLSSVAAPVKWCYGNVLRPGDKYSLRQEGAMLEL 1200

QY 3242 VVRNLRPODSGRYSCSFGDQTTSATLTVTALPAQFIGLKNKEATEGATATLRCLSKTA 3301
DB 1201 VVRNLRPODSGRYSCSFGDQTTSATLTVTALPAQFIGLKNKEATEGATATLRCLSKTA 1260

QY 3302 PVWRKGSSETLRDGDYCLRDQAGMCELOIRGLAMVDAAEYSCVCGEERTSASLTIRPMP 3361
DB 1261 PVWRKGSSETLRDGDYCLRDQAGMCELOIRGLAMVDAAEYSCVCGEERTSASLTIRPMP 1320

QY 3362 AHFTGLRLHQBSIEGATATLRCLSKAAAPVWRKGRESLRDGDHSLRDQAGVCELOICG 3421
DB 1321 AHFTGLRLHQBSIEGATATLRCLSKAAAPVWRKGRESLRDGDHSLRDQAGVCELOICG 1380

QY 3422 LAVADAGEYSCVCGEERTSATLTVKALPAKTEGLRNEEAVEGATAMLWCELS KVAPEVW 3481
DB 1381 LAVADAGEYSCVCGEERTSATLTVKALPAKTEGLRNEEAVEGATAMLWCELS KVAPEVW 1440

QY 3482 RKGPNLRDGDYILRQSGTRCELOICGLAMADAGEYLCVCGEERTSATLTIRALPARFI 3541
DB 1441 RKGPNLRDGDYILRQSGTRCELOICGLAMADAGEYLCVCGEERTSATLTIRALPARFI 1500

QY 3542 EDVKNQEAAREGATAVLOCELNSAAPVWRKGSSETLRDGDYSLRDGDKCELOIRGLAMA 3601
DB 1501 EDVKNQEAAREGATAVLOCELNSAAPVWRKGSSETLRDGDYSLRDGDKCELOIRGLAMA 1560

QY 3602 DTGYSVCVCGEERTSATLTVRALPIKFTTEGLRNEEAEGATAVLRCELS KNAPEVWVKGH 3661
DB 1561 DTGYSVCVCGEERTSATLTVRALPIKFTTEGLRNEEAEGATAVLRCELS KNAPEVWVKGH 1620

QY 3662 ETLRDGDHSLRDQAGARCELOIRGLVADAGEYLCMCKERTSAMLTVRAMPKSFIEGLR 3721
DB 1621 ETLRDGDHSLRDQAGARCELOIRGLVADAGEYLCMCKERTSAMLTVRAMPKSFIEGLR 1680

QY 3722 NEEATEGDTATLWCELS KAAAPVWRKGHETLRDGDHSLRDGSGRCELOIRGLAVVDAGE 3781
DB 1681 NEEATEGDTATLWCELS KAAAPVWRKGHETLRDGDHSLRDGSGRCELOIRGLAVVDAGE 1740

QY 3782 YSCVCGEERTSATLTVRALPARFIEDVKNQEAAREGATAVLOCELNS KAAAPVWRKGSSETLR 3841
DB 1741 YSCVCGEERTSATLTVRALPARFIEDVKNQEAAREGATAVLOCELNS KAAAPVWRKGSSETLR 1800

QY 3842 GGDYSLRDQTRCELOIRGLSVADTGYSCVCGEERTSATLTVRAPQVFPREPQOSLOA 3901
DB 1801 GGDYSLRDQTRCELOIRGLSVADTGYSCVCGEERTSATLTVRAPQVFPREPQOSLOA 1860

QY 3902 EEGSTATLQCELSBPTATVWVSKGLOLQANRRPREPLOGGTAEVLVLOLQREDGEYTC 3961
DB 1861 EEGSTATLQCELSBPTATVWVSKGLOLQANRRPREPLOGGTAEVLVLOLQREDGEYTC 1920

Db 842 WDRAPTFLRELSDETVVLQGVSTLACQVSAQPAQAATWSKDGPALSSSRVLISATLKNF 901
QY 7519 QLLTILVVAEDLGVVTCVSNALGVTTTGVLRKAERSSPCPDIGVYADGVLLVWK 7578
Db 902 QLLTILVVAEDLGVVTCVSNALGVTTTGVLRKAERSSPCPDIGVYADGVLLVWK 961
QY 7579 PVESYGPVTYIYQCSLEGGSWTTLASDIFDCCVLTSLKSRGGTYTFRACVSKAGMGPS 7638
Db 962 PVESYGPVTYIYQCSLEGGSWTTLASDIFDCCVLTSLKSRGGTYTFRACVSKAGMGPS 1021
QY 7639 SPSEQVLLGGPSHLASEESQGRSAQPLSTTKTFAPQTQIQGRFSVVRQWKEKASGRAL 7698
Db 1022 SPSEQVLLGGPSHLASEESQGRSAQPLSTTKTFAPQTQIQGRFSVVRQWKEKASGRAL 1081
QY 7699 AAKITPYHPKDTAVLREVEALKGRHPLAOLHAAYLSPRHLVILELCSGPELLPCL 7758
Db 1082 AAKITPYHPKDTAVLREVEALKGRHPLAOLHAAYLSPRHLVILELCSGPELLPCL 1141
QY 7759 ERASYESEVKDYLMQMLSATQYLHNQHLHLDLSENMIITEYNLLKVVDLGNQASLSQ 7818
Db 1142 ERASYESEVKDYLMQMLSATQYLHNQHLHLDLSENMIITEYNLLKVVDLGNQASLSQ 1201
QY 7819 EKVLPSDKPKVLETWAPELLGGQAVPOTDIWAIGVTAFIMLSAEYPVSSSEGARDLORG 7878
Db 1202 EKVLPSDKPKVLETWAPELLGGQAVPOTDIWAIGVTAFIMLSAEYPVSSSEGARDLORG 1261
QY 7879 LRKGLVRLSRCYAGLSGGAVAFRLSTLCAQWGRPCASSCLOCPWLTEGPAACSRPAPVT 7938
Db 1262 LRKGLVRLSRCYAGLSGGAVAFRLSTLCAQWGRPCASSCLOCPWLTEGPAACSRPAPVT 1321
QY 7939 FPTARLVFVRNREKRALLYKRHNLAQVR 7968
Db 1322 FPTARLVFVRNREKRALLYKRHNLAQVR 1351

RESULT 13
ADJ69186
ID ADJ69186 standard; protein; 1596 AA.
AC ADJ69186;
XX
DT 06-MAY-2004 (first entry)
XX Human heat mitochondrial protein as a therapeutic target SeqID992.
DE
KW mitochondrial; human; screening assay; diabetes mellitus;
KW Huntington's disease; osteoarthritis;
KW Leber's hereditary optic neuropathy; LHON;
KW mitochondrial encephalopathy lactic acidosis and stroke; MELAS;
KW myoclonic epilepsy ragged red fibre syndrome; MERRF; cancer;
KW neuroprotective; nootropic; antidiabetic; anticonvulsant; antiarthritic;
KW osteopathic; ophthalmological; cytostatic.
XX

OS Homo sapiens.
XX
XX WO2003087768-A2.
XX
XX 23-OCT-2003.
XX
XX 04-APR-2003; 2003WO-US010870.
XX
XX 12-APR-2002; 2002US-0372843P.
XX 17-JUN-2002; 2002US-0389987P.
XX 20-SEP-2002; 2002US-0412418P.
XX
XX (MITO-) MITOKOR.
PA
PA (BUCK-) BUCK INST AGE RES.
XX
XX Ghosh SS, Fahy ED, Zhang B, Gibson BW, Taylor SW, Glenn GM;
PI Warnock DE;
XX WPI; 2003-845369/78.

XX Identifying a mitochondrial target for drug screening assays and for
PT treating diseases associated with altered mitochondrial function,
PT comprises detecting a modified polypeptide in a sample and correlating
XX with the disease.
PS Claim 1; SEQ ID NO 992; 180pp; English.
XX
CC This invention relates to novel mitochondrial targets that can be used
CC for therapeutic intervention in treating a disease associated with
CC altered mitochondrial function. Specifically, it refers to a method for
CC identifying proteins of the human heart mitochondrial proteome that are
CC useful for drug screening assays, as well as therapeutic targets. The
CC present invention describes a method for identifying such proteins that
CC can be used in the treatment of various diseases associated with altered
CC mitochondrial function including diabetes mellitus, Huntington's disease,
CC osteoarthritis, Leber's hereditary optic neuropathy (LHON), mitochondrial
CC encephalopathy lactic acidosis and stroke (MELAS), myoclonic epilepsy
CC ragged red fibre syndrome (MERRF) or cancer. Accordingly, these
CC compositions have neuroprotective, nootropic, antidiabetic,
CC anticonvulsant, antiarthritic, osteopathic, ophthalmological and
CC cytostatic activities. This polypeptide sequence is a human heart
CC mitochondrial protein of the invention.
XX
SQ Sequence 1596 AA;

Query Match 16.0%; Score 6619.5; DB 7; Length 1596;
Best Local Similarity 81.7%; Pred. No. 0;
Matches 1309; Conservative 70; Mismatches 203; Indels 21; Gaps 7;
QY 2726 RARVRHDLHVGIITKRLKTMVLESGSCFECVLSHESASDPAMWTGGKTVGSSSRFQA 2785
Db 1 RARVRHDLHVGIITKRLKTMVLESGSCFECVLSHESASDPAMWTGGKTVGSSSRFQA 60
QY 2786 TRQGRKYLIVVREAAAPSDAGEVWFSVRLGTSKASLIIVRERPAALIKPLEDQWAPGEDVE 2845
Db 61 TRQGRKYLIVVREAAAPSDAGEVWFSVRLGTSKASLIIVRERPAALIKPLEDQWAPGEDVE 120
QY 2846 LRCLSRAGTPVHLKDKRAIRKSKQYDVVCEGTMAMLVIRGASLKDAGEYTCVEASKS 2905
Db 121 LRCLSRAGTPVHLKDKRAIRKSKQYDVVCEGTMAMLVIRGASLKDAGEYTCVEASKS 180
QY 2906 TASLHVEKANCFTTELTNLQVEEKGTAFTCKTEHPAAVTVMRKGLELRASGKHQPSQ 2965
Db 181 TASLHVEKANCFTTELTNLQVEEKGTAFTCKTEHPAAVTVMRKGLELRASGKHQPSQ 240
QY 2966 EGLTLRLTISALEKADSDTYTCDIGQASRAQLLVQGRRVHIIEDLEDVDVQEGSSATFR 3025
Db 241 EGLTLRLTISALEKADSDTYTCDIGQASRAQLLVQGRRVHIIEDLEDVDVQEGSSATFR 300
QY 3026 CRISPANVPEVHWFLDKTPHLANELNEIDAQPGGYHVLTLRQLAKDSGTIYFEAGDORA 3085
Db 301 CRISPANVPEVHWFLDKTPHLANELNEIDAQPGGYHVLTLRQLAKDSGTIYFEAGDORA 360
QY 3086 SAALRVTEKPSVFSRELTDAITTEGEDTLVCETSTCDIIPMCWTGDKGTLRGSARCOLSH 3145
Db 361 SAALRVTEKPSVFSRELTDAITTEGEDTLVCETSTCDIIPVCWTGDKGTLRGSARCOLSH 420
QY 3146 EGHRAQLLITGATLQDSGRYKCEAGGACSSIVRVHARPVRFQALKDLVLEGAATLR 3205
Db 421 EGHRAQLLITGATLQDSGRYKCEAGGACSSIVRVHARPVRFQALKDLVLEGAATLR 480
QY 3206 CVLSSVAAPVKWCYGNVLRPGDKYSLRQEGAMELVVRNLRPODSGRYSCSFGDQTTSA 3265
Db 481 CVLSSVAAPVKWCYGNVLRPGDKYSLRQEGAMELVVRNLRPODSGRYSCSFGDQTTSA 540
QY 3266 TLTVTALPAQFIGLKRNEKATEGATATLRCELSKTAPEVWRKSETLRDGDYCLRDGA 3325
Db 541 TLTVTALPAQFIGLKRNEKATEGATATLRCELSKTAPEVWRKSETLRDGDYCLRDGA 600
QY 3326 MCELQIRGLAWDAAEYSCVCGERTSASLTIRMPAHFTIGRLRHQESI EGATATLRCEL 3385
Db 601 MCELQIRGLAWDAAEYSCVCGERTSASLTIRMPAHFTIGRLRHQESI EGATATLRCEL 660


```
QY 3386 SKAAPVWVRKGRSLRDGDRHSRLQDGAVALCELOICGLAVADAGEYSCVCGERTSATITV 3445
D 661 SKAAPVWVRKGRSLRDGDRHSRLQDGAVALCELOICGLAVADAGEYSCVCGERTSATITV 720
QY 3446 KALPAKFTGLRNEEAVEGATAMLCWELSKVAPVWVRKGPENLRDGDRIYLRQEGTRCEL 3505
D 721 KALPAKFTGLRNEEAVEGATAMLCWELSKVAPVWVRKGPENLRDGDRIYLRQEGTRCEL 780
QY 3506 QICGLAMADAGBYLCVCGQERTSATITRALPARFIEDVKNQEARREGATAVLQCELSAA 3565
D 781 QICGLAMADAGBYLCVCGQERTSATITRALPARFIEDVKNQEARREGATAVLQCELSAA 840
QY 3566 PVWVRKGSSETLRDGRYSRLQDGTCKELQIRGLAMADTGEYSCVCGQERTSATITRALP 3625
D 841 PVWVRKGSSETLRDGRYSRLQDGTCKELQIRGLAMADTGEYSCVCGQERTSATITRALP 900
QY 3626 IKFTGLRNEEATEGATAVLRCELKMAVWVRKGPENLRDGDRIYLRQEGTRCELQIRG 3685
D 901 IKFTGLRNEEATEGATAVLRCELKMAVWVRKGPENLRDGDRIYLRQEGTRCELQIRG 960
QY 3686 LVAEDAGBYLCVCGQERTSATITRALPARFIEDVKNQEARREGATAVLQCELSAA 3745
D 961 LVAEDAGBYLCVCGQERTSATITRALPARFIEDVKNQEARREGATAVLQCELSAA 1020
QY 3746 RKGHETLRDGRHSRLQDGRSCELQIRGLAVDAGEYSCVCGQERTSATITRALPARP 3805
D 1021 RKGHETLRDGRHSRLQDGRSCELQIRGLAVDAGEYSCVCGQERTSATITRALPARP 1080
QY 3806 EDVKNQEARREGATAVLQCELSAAVWVRKGSSETLRDGRYSRLQDGTCKELQIRGLSVA 3865
D 1081 EDVKNQEARREGATAVLQCELSAAVWVRKGSSETLRDGRYSRLQDGTCKELQIRGLSVA 1140
QY 3866 DTGEYSCVCGQERTSATITRALPARFIEDVKNQEARREGATAVLQCELSAA 3925
D 1141 DTGEYSCVCGQERTSATITRALPARFIEDVKNQEARREGATAVLQCELSAA 1199
QY 3926 GLOQANGRRPRLOGCTAELVLDIQRDGTGEYTCGSOATSATITVTAAPVRLREL 3985
D 1200 PETLRDGRYSRLQDGTCKELQIRGLSVA 1259
QY 3986 OHQEVDEGTATLCELSELGRAGASVWVRKGSSETLRDGRYSRLQDGTCKELQIRGLSVA 4045
D 1260 RNEEAMEGATATLQCELSAAVWVRKGSSETLRDGRYSRLQDGTCKELQIRGLSVA 1318
QY 4046 GDYTDGTHTOSMASLSVRPKFKTRQLQSLQEGTGTDIARLCCQLSDAESGAVVQWLKE 4105
D 1319 GYVSCVCGQERTSATITRALPARFIEDVKNQEARREGATAVLQCELSAAVWVRKGS 1375
QY 4106 GVHLHAGPKYMRSGATRELLIHOLEAKDTGEYACVTGGQKTAASLRVTEPEVTIVRGL 4165
D 1376 PNTLRDGRYSRLQDGTCKELQIRGLSVA 1435
QY 4166 VDAEVTADDEVEFSCEVRAGATGVQWCLQGLPLOSNEVTEVAVRDRGRIHTLRKGVTE 4225
D 1436 RNEEAMEGATATLQCELSAAVWVRKGSSETLRDGRYSRLQDGTCKELQIRGLSVA 1492
QY 4226 DAGTVSFHLGNHASSAQITVRAPEVTILEPLQDVLQSEGDQASFCRLSRAGQARWAL 4285
D 1493 DTGEYLCVCGQERTSATITRALPARFIEDVKNQEARREGATAVLQCELSAAVWVRK 1550
QY 4286 GGVPLQANEMNDITVEQGTLLHLTLKVTLEDAQTVSFHVTG 4328
D 1551 GPETLRDGRHSRLQEN-----RLNPGGCGGCGELGSC 1582
```

RESULT 14

AAB30570

ID AAB30570 standard; protein; 871 AA.

XX AAB30570;

AC AAB30570;

XX DT 19-MAR-2001 (first entry)

XX DE A splice variant of a signal transduction polypeptide.

XX KW Signal transduction; H19G5; kinase; cardiac disease; angina pectoris; congestive heart failure; dilated congestive cardiomyopathy; hypertrophic cardiomyopathy; restrictive cardiomyopathy; hypertension; mitral valve disease; aortic valve disease; tricuspid valve disease; myocardial infarction; cardiac arrhythmia; arteriosclerosis; atherosclerosis; cardiac tumour; microbial infection; splice variant.

OS Homo sapiens.

XX WO2000063381-A1.

XX 26-OCT-2000.

PF 11-APR-2000; 2000WO-US009488.

PR 16-APR-1999; 99US-0129553P.

XX (SCIO-) SCIOS INC.

XX Zeng W, Stanton L, Kong H;

PI WPI; 2001-007013/01.

DR Novel h19G5 polypeptides capable of regulating signal transduction and exhibiting kinase activity useful for identifying antibodies to treat cardiac diseases, and additional mediators of signal transduction.

XX Claim 1; Page 74-76; 81pp; English.

CC The present sequence represents a splice variant of human in signal transduction polypeptide. The polypeptide is designated H19G5. The protein is capable of regulating signal transduction and exhibits kinase activity. The H19G5 transcript is expressed in the heart. H19G5 polypeptides and polynucleotides are useful for preventing or treating a cardiac disease, such as congestive heart failure, dilated congestive cardiomyopathy, hypertrophic cardiomyopathy, restrictive cardiomyopathy, mitral valve disease, aortic valve disease or tricuspid valve disease, angina pectoris, myocardial infarction, cardiac arrhythmia, pulmonary, arterial or renovascular hypertension, arteriosclerosis, atherosclerosis and cardiac tumours in humans. The polypeptide is also useful for detecting the expression of a protein capable of regulating signal transduction or the expression of a protein capable of acting as a donor or acceptor molecule of a phosphate group. The monoclonal antibodies can be used as probes for detecting discrete antigens expressed by tissue or cell samples, and therefore used in humans for localization and monitoring of microbial infection.

XX Sequence 871 AA;

Query Match 11.0%; Score 4533; DB 4; Length 871;

Best Local Similarity 100.0%; Pred. No. 3.8e-226; Mismatches 0; Indels 0; Gaps 0;

Matches 871; Conservative 0;

QY 7098 MGPDISLUPGRPKGPCSSPGSSAQSSQSSSLRVGSSQVGTPEGPSLDAEGWTQAEAD 7157

D 1 MGPDISLUPGRPKGPCSSPGSSAQSSQSSSLRVGSSQVGTPEGPSLDAEGWTQAEAD 60

QY 7158 LSDSTPTLQRPQEQATMRKFSLLGGGGYAGVAGYCTFAFGDAGMLGGPMMARIAWAV 7217

D 61 LSDSTPTLQRPQEQATMRKFSLLGGGGYAGVAGYCTFAFGDAGMLGGPMMARIAWAV 120

QY 7218 SQSEEEQEEARAESQSEEQEEARAESPLPVASRPVPEVGRAPTRSPPTPWEIDIGV 7277

D 121 SQSEEEQEEARAESQSEEQEEARAESPLPVASRPVPEVGRAPTRSPPTPWEIDIGV 180

QY 7278 SLVQIRDLSDGDAEADTISLDISEVDPAYLNLSDLYDKILPFFPMIFRKVPKSAQPEPP 7337

D 181 SLVQIRDLSDGDAEADTISLDISEVDPAYLNLSDLYDKILPFFPMIFRKVPKSAQPEPP 240

QY 7338 SPMAEEELAEPEPTWPGELGPHAGLEITEESEDVALLAEAAVGRKRWKSSRSRSLP 7397

Db 241 SPWABEELAEFPPTWPGELGPHAGLEITESEEDVDALLAAAYGRKRKWSPPSRSLF 300
QY 7398 HFFGRHLPLDEPAELGLRERVKASVEHRIKGRPEGLEKEGPPRRKPKGLASFRLSGLK 7457
Db 301 HFFGRHLPLDEPAELGLRERVKASVEHRIKGRPEGLEKEGPPRRKPKGLASFRLSGLK 360
QY 7458 SWDRAPTFRLSDETVVLGQSVTLACQVSAOPAAQATWSKOGAPLESSSRVLISATLKN 7517
Db 361 SWDRAPTFRLSDETVVLGQSVTLACQVSAOPAAQATWSKOGAPLESSSRVLISATLKN 420
QY 7518 FOLLTLVVVAEDLGVYTCVSNALGTVTGTVLKAERPSSPCPDIGENVADGVLLVW 7577
Db 421 FOLLTLVVVAEDLGVYTCVSNALGTVTGTVLKAERPSSPCPDIGENVADGVLLVW 480
QY 7578 KPVESEGYPTIIVQCSLEGSSWTTLASDIFDCCYLTSLKSRGTYTFRACYSKAGMGYP 7637
Db 481 KPVESEGYPTIIVQCSLEGSSWTTLASDIFDCCYLTSLKSRGTYTFRACYSKAGMGYP 540
QY 7638 SGPSEQVLLGGPSHLASEBESQGRSAQPLPSTKTFAQTQIORGRFVVVRQCEKASGRA 7697
Db 541 SGPSEQVLLGGPSHLASEBESQGRSAQPLPSTKTFAQTQIORGRFVVVRQCEKASGRA 600
QY 7698 LAAKIIPYHPKDKTAVLREVEALKGLRHPHQAOLHAAYLSPRHLVLIILELCSGPPELLCL 7757
Db 601 LAAKIIPYHPKDKTAVLREVEALKGLRHPHQAOLHAAYLSPRHLVLIILELCSGPPELLCL 660
QY 7758 ABRASYSESEVKDYLMQWLSATQYLHNQHLHLDRSENMIITEYNLLKLVVDLGNQAQSL 7817
Db 661 ABRASYSESEVKDYLMQWLSATQYLHNQHLHLDRSENMIITEYNLLKLVVDLGNQAQSL 720
QY 7818 QSKVLPDSKFDKYLETMAPLELGGQGVQPTDIWAGVTAFIMLSAEYVPSSEGARDLQR 7877
Db 721 QSKVLPDSKFDKYLETMAPLELGGQGVQPTDIWAGVTAFIMLSAEYVPSSEGARDLQR 780
QY 7878 GLRKGVLRLSRVAGLSGAVAFRLSTLCAQPMWGPCASSCLOCPLWTEEGPACSRPAPV 7937
Db 781 GLRKGVLRLSRVAGLSGAVAFRLSTLCAQPMWGPCASSCLOCPLWTEEGPACSRPAPV 840
QY 7938 TPTPTARLRFVVRNREKRALLYKRHNLAQVR 7968
Db 841 TPTPTARLRFVVRNREKRALLYKRHNLAQVR 871

RESULT 15

ID AAE16274
AC AAE16274 standard; protein; 871 AA.

XX AAE16274;

XX 26-MAR-2002 (first entry)

XX Human kinase PKIN-20 protein.

XX Human; kinase; PKIN-20; cancer; leukaemia; adenocarcinoma; osteoporosis;
XX immune disorder; atherosclerosis; Crohn's disease; Hodgkin's disease;
XX Acquired Immune Deficiency Syndrome; AIDS; Addison's disease; anaemia;
XX allergy; asthma; adult respiratory distress syndrome; multiple sclerosis;
XX autoimmune thyroiditis; bronchitis; diabetes mellitus; osteoarthritis;
XX Good pasture's syndrome; Graves' disease; pancreatitis; psoriasis;
XX rheumatoid arthritis; ulcerative colitis; cirrhosis; Cushing's syndrome;
XX hepatitis; hypothyroidism; cerebral palsy; cataract; angina pectoris;
XX cardiovascular disease; hypertension; vasculitis; myocarditis; obesity;
XX congestive heart failure; ischaemic heart disease; lung tumour; gout;
XX fatty liver; Niemann-Pick's disease; gene therapy.

OS Homo sapiens.

XX Key Location/Qualifiers
XX Domain 575..827
XX Domain /note= "Eukaryotic protein kinase domain"
XX Domain 580..812
XX Domain /label= Protein_kinase_domain

XX WO200196547-A2.
XX 20-DEC-2001.
XX 14-JUN-2001; 2001WO-US019444.
XX 15-JUN-2000; 2000US-0212073P.
XX 23-JUN-2000; 2000US-0213467P.
XX 30-JUN-2000; 2000US-0215651P.
XX 07-JUL-2000; 2000US-0216605P.
XX 13-JUL-2000; 2000US-0218372P.
XX 25-AUG-2000; 2000US-0228056P.
XX (INCY-) INCYTE GENOMICS INC.
XX Yue H, Lal P, Bandman O, Borowsky ML, Au-Young J, Lu Y;
XX Gandhi AR, Tribouley CM, Wallia NK, Lu DM, Greenwald SR;
XX Ramkumar J, Griffin JA, Kearney L, Burford N, Nguyen DB, Tang YT;
XX Baughn MR, He A, Thornton M, Hafalia A, Patterson C, Gururajan R;
XX Lo TP, Khan F, Recipon SA, Azimzai Y, Policky JL, Ding L;
XX Grether M, Elliott VS, Thangavelu K, Batra S, Ison CH;
XX WPI; 2002-090207/12.
XX N-PSDB; AAD26467.
XX New polypeptides, useful for diagnosing, treating or preventing disorders
XX of growth and development, cardiovascular and lipid, and diseases such as
XX cancer, comprise human kinase polypeptides.
XX Claim 1; Page 164-165; 197pp; English.
XX The invention relates to human kinase PKIN proteins and their
XX corresponding cDNAs. A composition containing PKIN agonist is useful for
XX treating a disease or condition associated with decreased expression of
XX PKIN and a composition comprising PKIN antagonist is useful for treating
XX a disease or condition associated with overexpression of PKIN. The
XX disorders include cancer (leukaemia, adenocarcinoma, lymphoma, melanoma,
XX myeloma, sarcoma, teratocarcinoma, Hodgkin's disease); immune disorder
XX (Acquired Immune Deficiency Syndrome (AIDS), asthma, Addison's disease,
XX atherosclerosis, anaemia, allergies, adult respiratory distress syndrome,
XX autoimmune thyroiditis, gout, bronchitis, Crohn's disease, diabetes
XX mellitus, multiple sclerosis, Good pasture's syndrome, Graves' disease,
XX osteoarthritis, osteoporosis, pancreatitis, psoriasis, Reiter's syndrome,
XX rheumatoid arthritis, Sjogren's syndrome, uveitis, ulcerative colitis,
XX bacterial, parasitic, fungal, viral, protozoal and helminthic infections)
XX growth and development disorders (arteriosclerosis, cirrhosis, hepatitis,
XX Cushing's syndrome, hypothyroidism, cerebral palsy, cataracts); cardio
XX vascular disease (arteriovenous fistula, hypertension, vasculitis,
XX aneurysms, congestive heart failure, angina pectoris, myocarditis,
XX ischaemic heart disease, chronic bronchitis, lung tumours); lipid
XX disorder (fatty liver, Fabry's disease, Niemann-Pick's disease,
XX hypocholesterolaemia, obesity). PKIN DNA is useful for assessing toxicity
XX of a test compound and in gene therapy. The present sequence is human
XX PKIN-20 protein

XX Sequence 871 AA;

Query Match 11.0%; Score 4529; DB 5; Length 871;
Best Local Similarity 99.9%; Pred. No. 6.2e-226;
Matches 870; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 7098 MGPDLSLPGRPKPGPCSPGSGASQSSQVSLRVSSQVGTGPGSLDAEGWTOBAED 7157
Db 1 MGPDLSLPGRPKPGPCSPGSGASQSSQVSLRVSSQVGTGPGSLDAEGWTOBAED 60
QY 7158 LSDSTPTLQRPQATMRKESLGRGGYAGVAGTGFAGCGDAGGMLGQCPMMARIWAV 7217
Db 61 LSDSTPTLQRPQATMRKESLGRGGYAGVAGTGFAGCGDAGGMLGQCPMMARIWAV 120
QY 7218 SQSEEEQEEARAESEEQEAEARASPLPQVSRPVPVEGRAPTRSSPPTPWEDIGQV 7277
Db 121 SQSEEEQEEARAESEEQEAEARASPLPQVSRPVPVEGRAPTRSSPPTPWEDIGQV 180

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6

Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 17:26:12 ; Search time 146 Seconds

(without alignments)
5251.063 Million cell updates/sec

Title: US-10-077-130-5

Perfect score: 41273

Sequence: 1 MDQPFSGAPFLTRPKAFV.....RNREKRALLYRHNLQVR 7968

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3577.5	8.7	26926	1 I38344	titin, cardiac mus
2	3313	8.0	7962	2 I38346	elastic titin - hu
3	2629	6.4	4162	2 T42633	connectin/titin -
4	2489	6.0	6642	2 T29757	protein UNC-89 - C
5	2048	5.0	5198	2 T43290	hemiscatin precurs
6	2043.5	5.0	5175	2 T20992	hypothetical prote
7	1636.5	4.0	7160	2 T27935	hypothetical prote
8	1634	4.0	6839	2 S57242	twitchin (similar
9	1619	3.9	6831	2 A88852	protein unc-22 (im
10	1392	3.4	6658	2 T13931	projectin - fruit
11	1074	2.6	4391	2 A38096	perlecan precursor
12	1046.5	2.5	6805	2 S20901	titin - rabbit (fr
13	1012	2.5	1906	1 S68235	myosin-light-chain
14	963.5	2.3	2783	2 T34416	hypothetical prote
15	869	2.1	1323	2 PN0568	connectin 3B - chi
16	785.5	1.9	3488	2 T34418	hypothetical prote
17	767	1.9	3707	2 S18252	heparan sulfate pr
18	703.5	1.7	1398	2 T25568	hypothetical prote
19	645.5	1.6	1176	2 JN0583	myosin-light-chain
20	622	1.5	1147	2 A59307	myosin-light-chain
21	571	1.4	1274	2 S55050	cardiac myosin-bin
22	532	1.3	1694	2 S50065	sialoadhesin - mou
23	528.5	1.3	1132	2 A35089	myosin-binding pro
24	520	1.3	1142	2 S36846	myosin-binding pro
25	519	1.3	3375	2 T19821	hypothetical prote
26	501.5	1.2	608	2 A35021	myosin-light-chain
27	480.5	1.2	610	2 A28798	myosin-light-chain
28	472	1.1	1896	2 T08851	Down syndrome cell
29	465.5	1.1	2541	2 T29340	hypothetical prote

30	463.5	1.1	451	2 S49128	twitchin-like prot
31	459	1.1	1612	2 T30805	ductin protein - mo
32	456.5	1.1	1398	2 T13741	hypothetical prote
33	456.5	1.1	1465	2 S43529	165K protein, skel
34	456	1.1	1423	1 I37275	death-associated p
35	456	1.1	1651	2 T14160	transmembrane rece
36	451.5	1.1	1051	1 JH0051	serine/threonine-a
37	451.5	1.1	1451	2 S42167	190K protein - hum
38	451	1.1	371	2 JC7733	death associated p
39	451	1.1	1123	2 S36846	myosin-binding pro
40	445	1.1	1734	2 A54602	microtubule-associ
41	441.5	1.1	374	1 S50193	Ca2+/calmodulin-de
42	439.5	1.1	1450	2 A44027	165K myofibrillar
43	427.5	1.0	1435	2 T32930	hypothetical prote
44	427	1.0	370	1 S57347	Ca2+/calmodulin-de
45	427	1.0	1138	2 S24614	myosin-binding pro

ALIGNMENTS

RESULT 1

I38344
titin, cardiac muscle [validated] - human
N:Alternate names: connectin
N:Contains: serine/threonine-specific protein kinase (EC 2.7.1.-)
C:Species: Homo sapiens (man)
C>Date: 12-Aug-1996 #sequence revision 12-Aug-1996 #text change 09-Jul-2004
C:Accession: I38344; I38345; S20898; S20899; S63665; S37393
R:Labelit: S.; Kolmerer, B.
Science 270, 233-296, 1995
A:Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
A:Reference number: A57430; MUID:96026330; PMID:7569978
A:Accession: I38344
A:Status: nucleic acid sequence not shown; translation not shown; translated from GB/EH
A:Molecule type: mRNA
A:Residues: 1-26926 <LAB1>
A:Cross-references: UNIPROT:Q10466; EMBL:X90568; NID:91017424; PID:91017425
R:Musco, G.; Tziatzios, C.; Schuck, P.; Pastore, A.
Biochemistry 34, 553-561, 1995
A:Title: Dissecting titin into its structural motifs: identification of an alpha-helix
A:Reference number: I38345; MUID:95119041; PMID:7819249
A:Accession: I38345
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1977-2014 <MUS>
A:Cross-references: EMBL:X83270; NID:G602579; PIDN:CAAS8243.1; PID:G602580
A:Note: conformation and properties are reported for a synthetic peptide corresponding
R:Labelit, S.; Gautel, M.; Lakey, A.; Trinick, J.
EMBO J. 11, 1711-1716, 1992
A:Title: Towards a molecular understanding of titin.
A:Reference number: S20897; MUID:92258380; PMID:1582406
A:Accession: S20898
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 13597-14200, I', 14202-14696 <LAB2>
A:Cross-references: EMBL:X64698; NID:G37192; PIDN:CAA45939.1; PID:G37193
A:Accession: S20897
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 16330-16382, S', 16384-16756, P', 16758-16860 <LAB3>
A:Cross-references: EMBL:X64699; NID:G37190; PIDN:CAA45940.1; PID:G37191
A:Accession: S20899
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 'P', 22278-22431, 'R', 22433-22448, 'G', 22450-22453, 'O', 22455-22480, 'TR', 22483-2

A:Cross-references: EMBL:X64697; NID:G37190; PIDN:CAA45938.1; PID:G37195
J. Mol. Biol. 256, 556-563, 1996
A:Title: Genomic organization of M line titin and its tissue-specific expression in two
A:Reference number: S63665; MUID:96177761; PMID:8604138
A:Accession: S63665
A:Status: nucleic acid sequence not shown

A:Molecule type: DNA
A:Residues: 26729-26825 <KOL>
A:Cross-references: EMBL:X92412; NID:g1236761
R:Gautel, M.; Leonard, K.; Labeit, S.
EMBO J. 12, 3827-3834, 1993
A:Title: Phosphorylation of KSP motifs in the C-terminal region of titin in differentiat
A:Reference number: S37393; MUID:94008990; PMID:8404852
A:Accession: S37393
A:Molecule type: mRNA
A:Residues: 26831-26926 <GAU>
R:Improta, S.; Politou, A.S.; Pastore, A.
submitted to the Brookhaven Protein Data Bank, February 1996
A:Reference number: A66736; PDB:1PIT
A:Contents: annotation; conformation by (1)H-NMR, residues 5253-5341
R:Fuhl, M.; Pastore, A.
submitted to the Brookhaven Protein Data Bank, August 1996
A:Reference number: A66201, PDB:INCT
A:Contents: annotation; conformation by (1)H-NMR, residues 'S', 26059-26155
C:Genetics:
A:Gene: GDB:TTN
A:Cross-references: GDB:127867; OMIM:188840
A:Map position: 2q31-2q32
C:Function:
A:Description: structural protein forming filaments in striated muscle
C:Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; pro
C:Keywords: alternative splicing; calmodulin binding; cardiac muscle; duplication; glyco
structural protein
F:24752-35008/Domain: protein kinase homology <KIN>
F:84,177,905,2276,2378,2459,2481,2563,2669,2763,2896,3088,3179,3384,3432,3628,3772,4068,
98,11066,11488,11515,11635,11949,12170,12478,12526,12645,12875,13001,13036,13295,13540,1
tatus predicted
F:16780,16976,17579,17602,17667,17681,17845,17899,18121,18188,18209,18336,18670,18680,18
21900,21935,22295,22495,22627,22897,23024,23318,23883,24012,24177,24290,24447,24642,248
F:26171,26178,26184,26190/Binding site: phosphate (Ser) (covalent) #status experimental
Query Match 8.7%; Score 3577.5; DB 1; Length 26926;
Best Local Similarity 20.2%; Pred. No. 3.6e-95;
Matches 1929; Conservative 1288; Mismatches 3662; Indels 2663; Gaps 320;
QY 4 POFGAPRLTRPKAFVSVGKDATLSQIVGNPTPOVSWEKDQOPVTAGARFLA---Q 60
Db 1030 PGEPAAPYFTRPVQKLVGGSVFVCGQVGNPXPVYVYKSGVPLTITGYKYVYNKQ 1089
QY 61 DGLVRLTLIDLALGDGQYVCRARNATGEAFAVGL-----QVD 100
Db 1090 TGEC-KLVISMTFADAGEYTIIVVRNKGHGETSASASLLEADYELLMKSQOEMLYQTQVT 1148
QY 101 A---EACACAEQAPFL----- 113
Db 1149 AFVQPEVGETAPGFYSEYEKEQALIRKKMAKDTVVVRTTVEBOEFHISFEERL 1208
QY 114 -----LRPTSIRVREGSEATFRGV 133
Db 1209 IKEIEYRIKTTLEELLEEDGEEKMAVDISESEAVESGFDLRKNYRILEGMYTFHCKM 1268
QY 134 GGSPPAVSWKDGRRLEPGDPRVVRBELGEASA-LRIRARPDDGTYEVRAENPLGA 192
Db 1269 SGYPLFKIAWKDGKRI--KHGERYQMDFLQDGRASLRIPVVLPEDEGIYTAFAASNIKGN 1326
QY 193 ASAAALVVDSDAADTASR--PGTSTAALLAHLQRR---REAMRAEGAPASPP----- 240
Db 1327 ALCSEKLVPEAPALGAPTYPTLEPVSKIRLSLSPRSVSRSPIRMSPARMSPARMSPARM 1386
QY 241 -----STGTR-----TCTVTTEGKHARLSYCVYTGPKPETVWKK 273
Db 1387 SPARMSPGRRLEETDESQLERLYKPVFLKPVFKLEGANCRCRDLKVVGRPMETFWFH 1446
QY 274 DGLVTEGRRH--VVVEDAQENFVLKILFCQSDRGLYTCTASNLVGGTYSSVLVV--- 328
Db 1447 DQQIVNDYTHKWKIKEDGTOSLI--IVPATPDSGSEWTVVAQNRAGRSSISVILTVEAV 1504
QY 329 ----- 328

Db 1505 EHOVKPMFVEKLNKXNVIKESGRLEMKVRATGNPNPDIVLKNKSDIIVPHKYPKIRIEGTK 1564
QY 329 -----R 329
Db 1565 GEALKIDSTVSQDSAWYATATANKAGRDTRCKNVNVEFAEPEPEKLLIPROTYRAK 1624
QY 330 EPAVP-----FKKLODLEVREKESATFLCS---VP 357
Db 1625 EIAAPELEPLHLRYQEQWEGBGLDYDKEQKQPFKKKLTSLRLKRFQPAHFCEKRLTIS 1684
QY 358 QPSTEAAWFKSETRI-----WASAKYGI-----BEEGTERLIT 390
Db 1685 DPTWVVEWLHDGKPLEANRLRMINEFGYCSLDGVAYSRDSGIITCRATNKYGDHTSA 1744
QY 391 VRNSADDDAVYICHTPEGRS-----TVASLAVQGNL-----LRKLPRKT 430
Db 1745 TLIVKDEKSLVESQLPEGRKGLQRIEELERNAHEGALTGVTTQKEKQKQKDPDILVYPPV 1804
QY 431 AVRVDGTAMF-CVELAVPVGPHWLRNQEYVAGRVAISAEGRTHLTITISQCCLEDVQG 489
Db 1805 RVLEGETARFCRVTPQPKVNVNLGQLIRKSRFRVRYDGI-HYLDIVDCKSYDTGE 1863
QY 490 VAFVAG-----DCQTSRFCVSAPRKPLQ-----PVD 518
Db 1864 KVTAENPEGVIEHKVKLEIQOQREDFRSLVRRAPEPEFHVHPEPKLQFEVQKVDREPD 1923
QY 519 -----PVVKARMESSVILSWSPPHGE-----RPVTIDGY-----LVEKKKLGTVTW 560
Db 1924 TTEYKEVVKLKRARIIEHKVPESEBELRSFKRTEGYVEAITAVELSKRKDESYEE 1983
QY 561 I-----RCHEAEWATPELTVAD-----VABEGNFQFRVSALNSFGQSPYLEPFGTVHLAPKL 613
Db 1984 LLRKTKDELLHW--TKELTEBKALAEKGIITTPFKPKDIELSPSME-----APKI 2034
QY 614 AVRPLKAVQAVEGEGVTFSVDLTVASAG-----EWFLDGQALKASS-VYEIHCDRTRH 666
Db 2035 PERIQSTQV--QGSDAHFRVR-----VGKPDPECEYKNGVKIERSDRIYVWPEDNVC 2088
QY 667 TLTRREVPASLHGAKLKF-----VANGIESSTRMEVRAAPGLTANKPPAAAAAREVLARLHE 722
Db 2089 ELVTRDTAB-DSASINVKAINIAGETSSHAFLVQAKLITFTQ-----ELQDVVAKEK 2143
QY 723 E-AOLLAELSDQAAVATWVKDGRTLSPGPKYEVQASAGRRLVLRDVARDDAGLYECV-- 779
Db 2144 TMAIFECETSEPFVKVWKYKDMGEVHEGDKYRMSDRKVHFLSLITLITDTSDAEYSCVLV 2203
QY 780 -SRGRIAYQLSVQGLARFLHKOMAGSCVDNAVAGGPAQFECETSEAHVHVHVKDGMELG 838
Db 2204 EDENVKTTAKLIVEGAVVEFVKELQD--IEVPESYSGELECTVSPENIEGKWHNDVELK 2261
QY 839 HSGERFLQEDVGTGTHRLVAATVTRQDEGTYSRVEDSDVFLRVSBPQVVFVFAKEQLARR 898
Db 2262 SNGKYTITSRG-RQNLTVKDVTKEDQGEYSFVIDGKTKCKLKM-KPRPAILIQLGSDQ 2319
QY 899 KLOBAAGASATLCEVAQAQTEVTWYKDGKLLSSSKVKCWEATGCTRRLLVQQAQADAG 958
Db 2320 KVC--EGDIVQLEVKVLSVESVEGWMKGQEVQPSDRVHVIDKQSHMLLIJEDMTKEDAG 2377
QY 959 EYS-----CEAGGQRLSFHLDVKEPKVVFVAKDQVAHSEVQAEAGANATLCEVAQAQ- 1010
Db 2378 NYSTPTPALGLSTGRVSVSDVITPL-----KDVNVIETGKVALCEKVSVPDV 2427
QY 1011 AEVNVYKDGKLLSSKLVHVEAKGRRRLVVQAGKTDAGDYSCEARGQVSRFLHITTEP 1070
Db 2428 TSVKWLNDLQIKPDDRVQAI VKGTQRLVINRTHASDEGPKYKLVGRVETNCLNLSVEKI 2487
QY 1071 XMPFAKEQSVHNEVQABAGASAMLSCEVAQAQTEVTWYKDGKLLSSSKVKCWEATGCTRR 1130
Db 2488 KIIRGLRDLTCTCTEQ-----NVVFEVLSHSGIDVLMNFKDKBIKPSKSKYKIEAHGKIYK 2542
QY 1131 LVLPQAGKADAGYSCEAGGQVSRFLHITPEKGVFAKEQSVHNEVQAEAGATTTAMLSCEV 1190
Db 2543 LTVLNMKDEGKTYTFAGENMTSGKL--TVAGASISK--PLTDQTVAES-QEAVFCEV 2597

Qy	1191	AQOTEVTWYKDGKGLSSSKVRMEVKGCTRRLLVVOQVGKADAGEYSCGAGQORVSFQLH	1250
Db	2598	ANPDSKGELWRDGHPLTNIRSESCHKRRLLIAATKLDODDIGEYTYKVATSKTSAKLK	2657
Qy	1251	ITEPAVFAKEQLVHNEVRTEAGASATLSCEVAQAQTE-VTWYKDGKGLSSSKVRLEAA	1309
Db	2658	V---EAVKIKKTL-KNUTVTET-ODAVFTVELTHPNVKGQWITKNGVVLSENEKIALSVK	2712
Qy	1310	GCMRQLVVOQAQADAGEYTCGAG--CQRLSFHLH----VSEPAVFAKEQLAHRKVQAE	1363
Db	2713	GTIYSLAIKNCIAVDESIVGYFRLGRLGASARLHVETVGIKKPKDVTALEN-----	2763
Qy	1364	AGAIALUSCEVAQAQ--TEVTWYKDGKGLSSSKVRMEAVGCTRRLLVVOQAQADTGEYS	1421
Db	2764	----ATVAFEVSVSHDTPVPKWPHKSVEIKPSDKHRLVSRKVKHKLMLQINISPDAGEYT	2819
Qy	1422	CEAGGQRLSFSLDVAEPKVPVFAKEOPVH----REVQAQAGASTLSCCEVAQAQTEVMWY	1476
Db	2820	AVNGOLECKAKLFV-----ETLHITKTWKNIEVPETKTASFCEVSHFNVSFWMWL	2869
Qy	1477	KDGKGLSFSSKVRMEAVGCTRRLLVVOQAQADAGEYSCAGSQRLSFPLHVAEPKAVPAK	1536
Db	2870	KNGVEIEMSEKFKIWWGKLHQLIIMNTSTEDSAEYTFVCGNDQVSAATLTVT-PIMITS-	2927
Qy	1537	EQPASREVQAEAGTSATLSCEVAQAQTEVTWYKDGKGLSSSKVRMEAVGCTRRLLVQEA	1596
Db	2928	---MLKDINAEKDTITFEVTVNYEGISYKWLKNGVEIKSTDCQMRTKKLTLSLIRNV	2984
Qy	1597	GQADAGEYSCAGDQRILSFHLHVAEPKVPFAKEQPAHREVQAEAGASATLSCEVAQAQTE	1656
Db	2985	HFGDAADYTFVAGKATSTALYVEARHIEPRKH---IKOIKVLKGRAMFECEVSEPDIT	3041
Qy	1657	VTWYKDGKGLSSSKVRMEAVGCTRRLLVVOQAQADAGEYSCGAGQRLSFRLHVAELEP	1716
Db	3042	VQWMDQDQELQITDRIKIQEKYVHRLIIPSTRMSDAGKYTVVAGGNVSTAKLFEVRDV	3101
Qy	1717	QISERPCEPREPLVYKHEHIIATLATPSNAVTWMLKDGVEI--RRSKHETASQGDTH	1774
Db	3102	RI--RS1KKEVQVLEKORAVV--EFEVNEDDVDHAWYKDGIEINFQVEHKKVYVERRIH	3157
Qy	1775	TLTVHGAQVLDSATISYCRVGAEGDPPVQVE-EVAAKFCRLBPVCBELGGTVTLACELS	1833
Db	3158	RMFISETRQSDAGEYTFVAGNRSSVTLYVNAPEPPQVLQELQPTVQSG-----K	3208
Qy	1834	PA--CAEVWRCGNTPRV-----GKRPQWVABGVRSLTVLGLURABDAGEYVC	1880
Db	3209	PARPFCAMI--SGRPQPKISWYKEEQLLSTGFCKKFLHDGOEYTLILLIEAFPDEDAAYTC	3265
Qy	1881	ESRDDH----TSAQLTVSYSPRVK-----FMSGLSTVVAEGCGEATQCVCVSPS	1925
Db	3266	EAKNDYGVATTSASLSVEFVPSVPOQMPVYPPIITPQDVTSSGQPARFCQVRSGT	3325
Qy	1926	DVAVWFRDGAALQPSKFAISQSGASHSLTISDLVLEDAGQITVEAGA---SSSAAL	1981
Db	3326	DLKVSWSYKDKIKPSRFRMTQEDFYQLEIAEAYPEDSGTYFVANNVAGQVSSTANL	3385
Qy	1982	RVREAPVLFKKULEPQ-----TVBERSVTLEVLTR-----	2013
Db	3386	SLEAPESILHERIEQETEMEMKBPSSFSFLSAEBGLHSAELQLSKINETILLELSSEPVPY	3445
Qy	2014	-----PPELRLWNTNATALAPGNVEI	2035
Db	3446	TKFDSKEBGTGPIPIKEVSNADISMGDVATLSVTVIGIPKPKIQWFFNGVLLTPSADYKF	3505
Qy	2036	HABGARHRLVLHNVGFADRGFFGCETPDD-----	2066
Db	3506	VFDGDDHSLILFTKLEDEGEYTCMASNDYKTCISAYLKINSKGEGHKDTETESAOKS	3565
Qy	2067	QAKL-----TVEMRQVRLVRGLQAV-EAREOGTATMEVQIASHADVCGSWTRDGLRQ	2117
Db	3566	LEKLGCPPHFLKELKPIRCAOGLPAIFEYTVVGEPAPTV-----TWFKENKQL-3615	

Qy	2118	QGPTC----	HLAVRGP--	MHTLTLSGLURP	PDSDGLMVF	KABGV----	HTSARLVV-----	2116																									
Db	3616	---CTS	VYTYIIHN	PNNGSGTF	VNDPQR	DSGLGYICK	AAEMLGESTCAAELVLLLEDDTD	3671																									
Qy	2162	-----	TELVSFSR-PL-	-----	QDVVTTEKE	KVTL	LECELSRPNVDVR	2198																									
Db	3672	MTDTP	CKAKSTP	PAPEDP	FPOTPLK	GAVALD	SEQEIATFVKOTILKAALITENQOOLS	3731																									
Qy	2199	-----	LKDGV	VELRAGK	MTAAAOQ----	ACRS	SLTYI-----	2225																									
Db	3732	EHI	KANELSSQLPL	GAELQ	SLILEQD	LTPESTR	ELFCINGSIHFPOLKEPSPNLQLOI	3791																									
Qy	2226	---RCE	PADOGVV	YVCD	ADAQS-----	-----	-----	2244																									
Db	3792	VQ	SOKTFS	KGILMP	EPETQ	AVLSDTEK	IFPSAMGTEQINSLTVEPLKTLAEPEGNYP	3851																									
Qy	2245	SAS	VYQGRTY-	-----	-----	-----	TLIYRRLAE-----	2265																									
Db	3952	QSS	IEPPMH	SILTS	VAEVL	SLKEKTV	SDTNRRQVRTLQKQEAQ	SALILSQSLAEGHVES	3911																								
Qy	2266	-----	-----	-----	-----	-----	-----	2287																									
Db	3912	LQSP	DMISQV	NEPLVP	SEHSC	TEGK	KILIESANPLENAGQDSAVRIEBSGLRPPLAL	3971																									
Qy	2288	-----	-----	-----	-----	-----	-----	2308																									
Db	3972	EEK	QVLLKEE	SDNVM	PPQDQI	IESKRE	PAIKVQEVQGRDLLSKESLLSGIPEEORLN	4031																									
Qy	2309	LEQ	VSRA-SAOV-	-----	-----	-----	-----	2322																									
Db	4032	LKI	QICRALQA	VAASQ	PLFSE	WLRIE	KEVVEAVNITQEPHIMCMYLVTS	SAKSVTEE	4091																								
Qy	2323	-----	-----	-----	-----	-----	-----	2327																									
Db	4092	VTII	IEDVD	POMANL	KWELR	DALCAI	IYEBIDILTAEGPRIQOGAKTSLQ	EBMDSFSGSQ	4151																								
Qy	2328	ELQ	GPKEYL	VSGL	YRKLII	SDVHA	EBEDETYTC	DAGV-----	2366																								
Db	4152	KVE	PITEPE	YESKY	LISTE	VSFNV	QSVQSRVQKY-LDATP	TVTKGVASAVVSEKQDES	LKPS	4210																							
Qy	2367	----KTS	AOFFVE	QOSITI	VRGLQ	DVTVM	EPAPA-----	-----	WPEC	TSLPSVR	2408																						
Db	4211	EEK	ESSSES	GBTE	AVATVKI	QEA	EGGLIKEDGPMIHTPLVD	TVSEEGDI	VHLTTSIT	NAK	4270																						
Qy	2409	PPK	WLLGKT	VLQAG	NVGL	BEQEG	VHRLMLRR-----	-----	TCSTMT	GPVHPTVGK	SR	2458																					
Db	4271	EVN	WYEN	KLVP	SDER	KFCLO	DQNTYTLVD	KVNTED	HQGEY	CEALDS-----	-----	GKTAT	4325																				
Qy	2459	SARL-V	VS	DI	PVVL	TRPLE	PKTGRE	LQSVVLS	CDFR	PAPKA-VQWY	KDDT	PLSP	EKFPM	2516																			
Db	4326	SAKL	TVVKRA	APV	IKRIE	PLE	VALGH	LAKFTCEI	QSA	PNVR	QFWPK	AGRIE	YESDKSI	4385																			
Qy	2517	SLG	QWAL	ILRL	MAD	AGVY	RCQAG	SAHS	TEVT	VEARE	VTVG-----	-----	PLQDA	2567																			
Db	4386	RSK	YI	SSLE	ILR	TVQV	DCGEY	TKAS	NEYG	SVSCTA-TL	TVTV	VPGEK	KVRKLL	PERKP	4444																		
Qy	2568	EAT	EGWAS	FSCEL	SHED	BEVENS	LN	GMPLY	NDSPFHEI	SHKRRH	TLVLKS	IORAD	AGIV	2627																			
Db	4445	EPKEE--	VVLKSV	LRRK	PEEBE	KVP	PKKL-----	-----	EKKV	PAPV	PEPPPP	KPVE	VEVE	PVT	4497																		
Qy	2628	RASS	LK	YSTS	AR	LEV	RKVPV	FLKAL	DDL	SAB	ERG	TLLAL	QCE	SDPE-----	-----	AHV	2678																
Db	4498	TKR	KI	PE	TKVP-EI	KPAI	PL	PA	PE	KPK	PAE	AVKTI	KPPV	PE	PEPT	PIA	APV	VPV	4556														
Qy	2679	WRK	DG	VOL	GPS	D	KVDFL	HTAG-TR	GL	VV	HD	V	SD	E	AG	LYT	CHVGS	-----	-----	EETR	ARVR	2731											
Db	4557	GK	AEAK--	APKE	AAK	PGI	KGV	PKTP	PSPIE	ARR	KL	RFG	SG	GEK	PPD	E	APFT	YQ	4613														
Qy	2732	HD	L	HVG	T	IKL	K	TM	VELE---	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	2787										
Db	4614	KAP	L	KV	KEI	KODI	L	IT	SE	F	V	G	SSA	I	F	CE	L	SP	STAI--	-----	-----	-----	TTWK	OG	NS	I	RE	SP	KH	RIA	4671		
Qy	2788	QGR	KY	IL	V	RE	AP	AS	D	AGE	V	VS	VR----	GLT	S	K	AS	L	I	V	R	P	AA	I	K	P	ED	Q--	W	V	AP	GE	2842


```
Db      8709 VVTRNKMCLNWSDPEDDGGSEITGFI--IERKAKM----- 8743
Qy      6516 LAALSHPLVTGLLDQFPETKRLILILELCSSEELDLRYKGVWTEAEVKVVIQQLVEGL 6575
Db      8744 -----HTWRQPIETERS-KCDITGLLEGQ 8766
Qy      6576 HYLHSHGVL-----HLDIKPSNLMVHPAREDIKICDFGAQNTTFAELQFSQYGS 6626
Db      8767 EX--KFRVIAKNKFGCGPPVEIGP--LLAVDP-----LGP 8797
Qy      6627 PRFVSPETIIQQNPVSEASDIWAGVISYLSLTCSPFAGESDRATL-LNVLEGRVSWSPG 6685
Db      8798 P--TSPE-----RLATV-----ERQSRITTLWDKEPRNGGSP 8828
Qy      6686 MAALHSE---DAKOFIKATLQAPQARPSAAQCLSHPWFLSKMAEAEAHFINTKQLKFL 6741
Db      8829 IGYITIEKRHRDKDPF-----ERVNKLCPSTTFLVENLDEHQMYEPRVK----- 8873
Qy      6742 LARSRWQSLMSYKSIILVNRSIPELLRGPPDSPSLGVARHLCRDTGCGSSSSSSSDNEIA 6801
Db      8874 -----AVNEIGSEPSLP-LNVV-----IQDDEVP 8897
Qy      6802 PPARAK-----SLP---PSPVTHSPF-----LHPRGFLRPSASL---PERAE 6837
Db      8898 PTIKRLSVRGDTIKVKAGEPVHIDPVTGLPMPKIEWSKNETVIEKPTDALQITKEEVS 8957
Qy      6838 ASERSTEAPAPPASPEGAGPPAAQCVPRHVSIRSLFYHQAGESPEHGALAPGSRHHPAR 6897
Db      8958 RSEAKTELISIPKAVREDKGTVTYVTSNRLGVSFRNVHVEYDR-----PSPPRNLAV 9009
Qy      6898 RRHLKGGYIAGALP---GLREPLMEHRVLEBEAAREEOA-----TLLAK-----A. 6940
Db      9010 TDIKAESCVLTWADPLDNGGSE--ITHYVIDKRDASRKKAEEWETNTAVEKRYGIWKL 9067
Qy      6941 PSFETALRLPAGTHLAPGHSHLSLHSDSPTRPSPSEACGEAQRLPSAPSGGAPTRDMGH 7000
Db      9068 PNGOYEFVRVAVNKY---GISDECKSDKVIQDP-----YRLPGPP--GKP-KVLAR 9113
Qy      7001 PGSKQLPST---GHPGHTAQ--PERPSPDSP-WGQAPPFCHPKQG-----SAPO--E 7045
Db      9114 TKGSLMVSWTPLDNGGSPITGYWLEKEEGSPYWSRVSRAPITKVGLKGVFNPRLL 9173
Qy      7046 GCS-PHPAVAPCPGSPFPFGSCKEAPLVSPSP-FLGQFQAPPAPAKASPPPLDSKMGPGDI 7103
Db      9174 GVKYQFRAMAINAAGIGPSE-----PSDDEVAGDPIFFPGP-PSCPVEVKDT-KSSI 9224
Qy      7104 SLPGPKPGPCSPGASQASSQVSSSLVGSQVGTGPGPSLDAEGH--TQEAEDLSDS 7161
Db      9225 SLGMRP---PAKDG---SPIKGYIVEMQEEGT-----TDWKRVPNEPKLITT 9266
Qy      7162 ---TPTLQRPQEQATMRKFSGLGRGGVAGVAGYCTFAFGDAGGMLGQGPWWARIAWAV 7217
Db      9267 CECVFNLUKE-----LKKYF-----RVK-AV 9287
Qy      7218 SQSEEEQEEARAEQSBEQFQARAEPLQVSARVPVEVG-----RAPTRS 7264
Db      9288 NEAGESEPSDTTGEIPATDIOEE-----PEVFI---DIGAQDCLVCKAGSQIRIPAVI 9337
Qy      7265 SPEPTP---WEDIGQVSLVQ---INDLSGDAEADTISLIDISEVDPAYNLNSLDYDI--K 7316
Db      9338 KGRPTPKSWEPDGRKAKMKDGVHDIPEDAQLETAENSSVIIIPCKRSHTGKYSITAK 9397
Qy      7317 YLPPFEMIFRKVPKSAQPEPPSPMAEELABFP-EPTWPMGELG----- 7360
Db      9398 NKAGQKTANCRVKVMVDPGPPKDLKVSIDITRGSCLSLWMDDDGCDRIKGYVIEKRTID 9457
Qy      7361 -----PHAGL-----EITEESEDVALLAAAVGRKRWKSWSPSRLSFHFPGRHLPL 7406
Db      9458 GKAWTKVNPDCGSTTFVVPDLLEQQYEFVRVRAENRFGIGPPVETIQTATTARDP--IYPP 9515
Qy      7407 DEPAELGLRERVKASV-----EHISRLIKGRPGELEKEG----- 7440
```

```
Db      9516 DPPIKLIKTLITKNTVHLNWKPPKNDGSPVTHYIVECLAWDPTGTCKEAWRQCNKRQVE 9575
Qy      7441 -----PPRKKXPGLASPRLSGLKSW-----DRAPTF-LRELSDETYYL 7476
Db      9576 ELQFTVEDLVGGEYEFVFKAVNAAGVSKPSATVGPCCQCPDMPDPSIDLKEFME--VEE 9633
Qy      7477 GQSVTLACQVSAQPAQAATWSKDGAPLESSRSRVLISATLKNFOLL--TILVVVAE---D 7530
Db      9634 GTNVNIVAKIKGVFPPTLTWFKAPPKPDNKEPVLVDTHVNKLVVDDTCTLVIPOSSRSD 9693
Qy      7531 LGVITCSVSNALGTVTGTVLRKAERPSSSPCPDIG-----EVYADGVLLVWKPVESVGP 7585
Db      9694 TGLYTTITAVNNLGTASKEMRLNVLGRPG-----PPVGPPIKFESVSADOMTSLWFFPKDDGG 9749
Qy      7586 ---VTYIVQ-CSLEGGSWTTIASDIFDCCYLTSLKSRGGTYTERTACVSKAGMG-PYSSP 7640
Db      9750 SKITNYVIEKREANRKTWVHVSSEPKECTYTIPLKEGHEVYFRIMAQNKGIGIEPLDSE 9809
Qy      7641 SE 7642
Db      9810 PE 9811

RESULT 2
138346
Elastic titin - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 09-Jul-2004
C:Accession: 138346
R:Labelit, S.; Kolmerer, B.
Science 270, 293-296, 1995
A:Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
A:Reference number: A57430; MUID:96026330; PMID:7569978
A:Accession: 138346
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-7962 <RES>
A:Cross-references: UNIPROT:Q10465; EMBL:X90569; NID:gi1017426; PID:gi101
C:Genetics:
A:Gene: GDB:TTN
A:Cross-references: GDB:127867; OMIM:188840
A:Map position: 2q31-2q31

Query Match      8.0%; Score 3313; DB 2; Length 7962;
Best Local Similarity 21.5%; Pred. No. 4.4e-88;
Matches 1710; Conservative 1105; Mismatches 3287; Indels 1868; Gaps 306;

Qy      10 PRFTTRPKAFVSVGKDATLSCQIVGNPTPOVSHVEKQQPVVTAGARFLAQDGLYRLTI 69
Db      10 PAITPLQDVTTSRGQPARFCRVSGTDL-KVSWYSKDKKIKPSRFFRMTQFEDTYQLEI 68
Qy      70 LDLALGDSGVVCRARNAIGEAFAAVGLQVDAEAACAEQ-----APHFLLRPT 117
Db      69 AEAPVEDEGTYTFVANNVAGQVSSSTANLSLSEAPSLIHETIEQEIEMEMKAAPIKRIE 128
Qy      118 SIRVREGSEATFCRVGSGSPRPAVSWSKDGRRLGPDGPRVVRBELGEASALRIARAAPR 177
Db      129 PLEVALGHAKFTCEIQSAPNVRQWFKAGRIEVESDKSIRSSKY--ISSLEILRTQVV 186
Qy      178 DGGTYEVRANPLGAASAAAALVVDSDAADTASRPGTSTAALLAHORRRREMARAEAGA 237
Db      187 DCGEYTCASNEYGSVSTATLTV-----TEA 213
Qy      238 SPSTGTGRTCTVTE--GKHARLSCVVTGEPKPEVWKDQGLVTEGRRHHVVEYDAQENFV 295
Db      214 YPPTFLSRPSLSTTFVGAAKAFICTVTGTVIETIQDGDGALSPSPNWR1-SORENKHI 272
Qy      296 LKILFCQSQDRGLTYCTASNLVG-QTYSVLLVVVREPAVPFKRLQDLVREKESATFLC 354
Db      273 LELSNLTIQDRGVYVSKASKNFGADICQAEILIIDKP--HFIKELEPVQSAINKKVHLEC 330
Qy      355 EVPOP-STEAAWFEETRLWASAKYIEEGTERLRLVRNVVSADDDAVYICETE---GS 410
```


Db 331 QVEDRKTVTWSDGQKLPCKDYKICPEDKIATLPLAKLKDSTGTYCTASNEAGSS 390
Qy 411 RTVAELAVQ--GNLLRKLPRKTAVRVGDTA-MFCVELAVPGVPVHWRNQBVEVAGGRAVA 467
Db 391 SCSATVTVPSPFVKYVLPGLPGESARLHCKLKGSPVQTWTFKNNKELSSNTVR 450
Qy 468 ISAEGRHTLTISQCCLEEDVGQAFMA---GDCOTSTRFCVSAPRKPLQPPVDPVKA 523
Db 451 MYFVNSEALDITDKVEDSGSYCEAVNDVSDSCSTEIVKEP--PSFIKLEPADIV 508
Qy 524 RMESVILSWSPPHGERPVITDGLVBKKGKLTWIRCHBAEWATPELTAVDAVEG 583
Db 509 R-----GTNALLQC-----EVSCTG 523
Qy 584 NQFRVSALNSPGQSPYLEFPCTVHLAPKLAVRTPLKAVQAVEGGEVTFSDVLTVASAGE 643
Db 524 PFEI-----S 528
Qy 644 WFLDQALKASSVYBIHCDTRHTLTIREVPASLHQAQLKFVANGIESSIRMEVRAAPCL 703
Db 529 WFKDKKQIRSSKYLRFQKLSVLCLLEIFSNADVGEYECVAVNEVGKGCMMATH-----583
Qy 704 TANKEPPAAAAAR--EYLABLHEBAQLLAEL-SDQAAAATWTKDGRTLSPGPKYEVOASAGR 760
Db 584 LLKEPPTFVKVDDLIALLGGQVTTLQAAVRGSEPISVTWKQOEVIREDGKIKMSFNGV 643
Qy 761 RVLLVRDARDAGLYECVSR---GGRIAY-OLSVQGLARFLHKDMAGSCVDVAVAGGPAQ 816
Db 644 AVLIIQVQISFGGKYTCCLAENEAGSQTSGBELIVKEPAKIER---AELIQVATGDPAT 700
Qy 817 PECETS-EAHVHVHVKDGMELGHGSEFLQEDVGTGRHLVAATVTRDEGTY----SCR 871
Db 701 LEYVAGTPELKPWKYDGRPLVAG-KKYRISFKNNVAQLKPYSAELHDSGGQTYFEISNE 759
Qy 872 VGEDSVDFRLRVSEPKV--FAKEQLARRKLOAEAGASATLSCEVA-QAOTEVTVYKDKG 928
Db 760 VGSSECEITFTVLDRIAPFFTK---PLRNVDSVNGTCRLDCKIAGSLPMRVSNFWDGK 816
Qy 929 KLSSSKVCMEATGCTRLVVOQAQADAGEVSCBA---GQRLSHFLDVKEPKVFP--982
Db 817 EIAASDRYIAFVEGTASLEIIRVDMNDAGNFTCRATNSVSGSKDSGALLIHOEPPSPFTK 876
Qy 983 --AKDOVAHSEVQAEAGANATLSCEVAQAAB---VMWYKDKKLSLKVHVEAKGR 1036
Db 877 PGSKDVLPGSAV-----CLKSTFGSTPLTIRWFKNGKELVSGGSCYITKEALE 925
Qy 1037 RRLVVOQAGKTDAGDYSC---ARGQVSRFLHITPEPKMMPAKQSVHNEVQAEAGASA 1092
Db 926 SSLELYLVKTSDSGYTCKVSNVAGVECSANLKVKEPATFVEK---LEPSQLLKKGDAT 982
Qy 1093 MLCSEV-AQAOTEVTVYKDKKLSLSSSKVGMVEKGCCTRLVLPOAGKADAGEYSCEAGGQ 1151
Db 983 QLACKVTGTPPIKITWIFANDREIKESSKHRMSFVESTAVLRITDVGIEDSGEYMCQAQNE 1042
Qy 1152 RVSEH---LHITEPKGVFAKEQSVHNEVQAEAGTAMLSCEVA-OPQTEVTVYKDKKLS 1207
Db 1043 AGSDHCSIVIKESPYTKE---FKPIEVLEKEYDVMLLAEVAGTPPFEITWFKDNTILR 1099
Qy 1208 SSSKVRMEVKGCTRLVVOQVKGADAGEYSC---EAGQVSRFOLHITPEKAVPAKEQL 1263
Db 1100 SGRKYKTFIQDLHLSLQILKFPVADAGHYQCVRTNEVSGSSICSAVRTIREPPSPFKK---1156
Qy 1264 VHNVRVTRBAGASATLSCEVAQAOTEVTVYKDKKLSLSSSKVRIEAGCMRQLVVOQAQOA 1323
Db 1157 IESTSSLRGGTAAFQATLKGLSPITVTVLWKSDEITEDDNRMTENNVAASLYLSGIEVK 1216
Qy 1324 DAGEYTC---EAGQVSRFOLHITPEKAVPAKEQLAHRKVAQAEAGATLSCEVAQAOT 1379
Db 1217 HDGKVCQAKNDAGIQRCSALLSVKEP-ATITEAVS---IDVTQGDPTALQVKSFGTK- 1271
Qy 1380 EWT--WYKDKKLSLSSSKVRMEVAGCTRLVVOQAQADAGEY-----SCEAGG 1426
Db 1272 EITAKWFKDQBELTIGSKYKISVTDVTSILKIIISTEKDSDGEYTFEVQNDVGRSSCKA-- 1329

Qy 1427 QRLSPSLDVABRPVVPFAKEQPVHREVOQAQAGASTTILSCEVAQAQ-TEVWYKDKKLSLSS 1485
Db 1330 -RINV-LDLIIIPP-SFTKKL---KQWDSIKGSFIDLECIVAGSHPIOSIOWFMDQOISAS 1383
Qy 1486 SKVRMEVAGCTRLVVOQAQADAGEYSC---EAGSQRSLPHLHVAEPKAVPAKEQOAS 1541
Db 1384 EYKFSFDHNTAFLEISQLEGTDSTYTCSATNKAGHQCOSGHLTVKEPP--YFVEKPOS 1441
Qy 1542 REVQAEAGTSATLSCEV-AQAOTEVTVYKDKKLSLSSSKVRMEVAGCTRLVVOQAQAD 1600
Db 1442 QDV--NPNTRVQLKALVGTAPMTIKFKDNKELHSGAARSVMKDDTSTLSLEFAAKATD 1499
Qy 1601 AGEYSCK---AGDORLSFHLHVAEPKVVFAKEQPAHREVOQAEGASATLSCEV-AQAOT 1655
Db 1500 SGTYICQLSNDVGTATSKATLTVKEPPQFIKKPSPV---LVLNRNQSTTFEQOITGTPKI 1556
Qy 1656 EVTVYKDKKLSLSSSKVRMEVAGCTRLVVOQAQADAGEYSC---AGGQRLSFLRHV 1711
Db 1557 RVSNWLDGNEITAQKHGISFIDGLATFQISGARVENSQVTVCEARNDAGTASCSELKV 1616
Qy 1712 ALEBPQISERPCRRREPLVVKHEHDIILATLATPSAATVTVLWKGVEIRRRKHRTASQ 1771
Db 1617 KEPPTFIRE---LKPVEVVKYSDVELECEVTGTPPEVTVLKNRREIRSSKKYTLTDRV 1672
Qy 1772 DTHLTIVHGAQVLSAIYSCRVAGBQDFPVQVEEVAAK---FCRLLEPVCBELGGTVT 1827
Db 1673 SVFNLIHTKDPSTGEYCQIVSNBEGGSCST-RVALKEPPSPFIKKIENTTTTVLKSSAT 1731
Qy 1828 LACEL--SPACAEVVMRCGNTQPRVGRKFMVAEPG---VRSULTVLGRAED---AG 1876
Db 1732 FQSTVAGSPPIIS-ITWL-----KDDQILDEDDNVYISFVDSVATLQIRSVONGHSG 1781
Qy 1877 EYVCESRD---DHTSAQLTVSVPRVVKFMSGLSTVVAEEGGEATFQCVVSPS-DVAVVM 1931
Db 1782 RYTQAKNESGVERCYAFLLVQEP--AQIVEKAKSDVTEKDPMTLECVVAGTPELKVKM 1839
Qy 1932 FRDGALLOPSKFPALSQSGASHLTIISDLVLEDAGQITVEAB---GASS-SNALRV--RE 1985
Db 1840 LKDGKQIVPSRYFSFMSFENNVSFRIQSVMKQDSQGYTFKVENDFGSSSCDAYLRLDQ 1899
Qy 1986 APVLFKKLEPQTVSEERSVTVLELTPWP-ELRWTNRNATALAPGKVEIHAEGARHL 2044
Db 1900 IPPSFTKKLTQMDKVLGSSIHMEKVSGLSIPISQWFKD-----GKEI---STSAKVL 1950
Qy 2045 VLNHVGFADRGFGCETPDDTKOAKLTVMERQVRLVRGLQAVAREQGTATMEVQLSHAD 2104
Db 1951 VCH-----ERSVSLVNNLEL-----EDTAN-----1971
Qy 2105 VDGSWTRDGLRFOQGPCHLAVRGPMTHTLTLSGLRPEDSGLMVFKAEGVHTSARLVVTEL 2164
Db 1972 -----YTCVKSNAVAGDDA-----CSGILTVKEP 1994
Qy 2165 PVSFSRP-LQDVV--TTEKEKVTLECELSRPNVDVRLWLDGVELBAGTKMTAAAGACRS 2221
Db 1995 PSFLVKPGRQAIIPDSTVEFKAILK---GTPPFKIKWFKDDDELVSQPKCFILGEGSTSF 2051
Qy 2222 LTIYRCEPADQGVVYCDADHAQSSASVKVQGRVTYTLIYRRVLAEDAGBIQFVAENAESRA 2281
Db 2052 LNLXSVDAKSTGQYTC-----VTNDVG-----SDSCTT 2080
Qy 2282 QLRVKELPVTILVRPLRDKIAMKRGVLECOVSRASAOVR--WFKGSELQBPQPKYELVS 2339
Db 2081 MLLVTEPKFVKKLEASKIVKAGDSRLECKIA-GSPBIRVWVFRNEHELSPASDYKRYMT- 2138
Qy 2340 DGLYRKLIISDVHAEDEDTYTCADGDVKTSAQFFVEEQSIITVRLQDVTWMEPAPAFE 2399
Db 2139 -----FTDSVAVTOMNVLSTEDSGD-----FI 2160
Qy 2400 CETSIPIVRPKLLGKTLVLAGGNVLEQBTGTVHRLMLRRTCTMTGPHVFTVKGSRSS 2459
Db 2161 CEANQPA-----GST-----SCST-----2174

Db 4263 TTGDTCTLECTVA-GTPELSTKWFPGDKEL-TSDNKYKISPNKYVGLKIINVA PSDSGV 4320
Qy 4504 VTFACRDV-----ASARUTVLGLPDPEDA-----EVAHSSHTVTLW 4543
Db 4321 YSFEVQNPVGDSCCTASLQVSDRTVPPSFRKLKETNGLSGSSVMECKVGVSPISVSW 4380
Qy 4544 AAPMSDGGGLCGYRVEKAGATGQWRLCHELVPCPECVVDGLAPGETYRFRVAAGVPV 4603
Db 4391 ---PHEGN-----EISSGRKYTTLTDNTC-----ALTVMLESDSDGYTCTATNAG 4426
Qy 4604 AGEVPHLPOTVRLAEPKPPVPPQAPESRQVAGEDVSLELEVVAEAG-EVIMHKGMR 4662
Db 4427 SDE-CSAFLTVR---EPPSFV---QKPDPMVLTGTNTFTTSIVKGTPPFSVSWFKGSSE 4479
Qy 4663 IQGGRFVSVQGRQOMLVIKFTAEDOGHYCHGLAOGSICPAATFOVALSPASVDEAP 4722
Db 4480 LVPDRCNVSLSDSVAEILELFDVDTSSQSEYTCIVS-----NEAG 4519
Qy 4723 QPSLPPEAAQEGDLHLLEALAR--KRMSREPTLDSISELPEEDRGRSRLPQEAEEVAP 4780
Db 4520 KASC-----TTHLYIKAKAFVVR-----LNDYSIEKGK-----P 4549
Qy 4781 DLSEGYSTADELARTGADLSHTSDDESRAGTPSLVTYLKKAGRPGTSPASKVGPAA 4840
Db 4550 LILEGTFT-----GTPPISVTWKXG----- 4570
Qy 4841 PSVPKQOOQEPAAVPPPLGDLSTKDLGDPMSDKAAVKIQAFKGYKVRKEMKQOEGPMF 4900
Db 4571 INVTPSORCNITTEKSPILSPSTVEDAGYNCYIE-NASGKD-SCSAQILILEPPVF 4628
Qy 4901 SHTTGDETAQVGDALRLFCVVASKADVRARMLKOGVEL---TDGRHHIDQLGDGTCULL 4957
Db 4629 VKQLEPVKSVGDSASLQCLAGTPEIGVSWYKGTDLRPTTYKQH-----FRNNVATLV 4684
Qy 4958 IAGLDADAGCYTCQVSNKFGQVTHSACVVVSGSESEASSGSGELDAFRRAARLRL 5017
Db 4685 FNQVDINDSGEYICKAENSVEVSASTFLTVOEQ-----KLPPSFSRQLRDVQE- 4733
Qy 5018 PRTPSPAEVDEELPLSADGEPABEPADQVTRDEBHFICIRPEALTEARQAVTRQOE 5077
Db 4734 -----TVGLPVVFDCAISG-----SEPIVSWSYKDG-----KPLKDSPNVQTSFLD 4774
Qy 5078 MFATLGI-GVEIKLVEQGRVEMCISKETAPVVPVPEPLSLLTSDAAPVFLTELQOE 5136
Db 4775 NTATLNIPTDRSLAGQ-----YSCTA-----TNPIGS--ASSARLILTEGKNPP 4818
Qy 5137 VQD-----GYPVSFDCVVTQOPMPSVRWFKDGLLEDDHYMINEDQOGHOLI 5186
Db 4819 FFDIRLAPVDVAVGESADFCHVTGTQPIKYSWAKDSREIRSGGKYQISYLENSAH-LTV 4877
Qy 5187 TAVVPADMGVYRCLAENSMGVSSTKAELRVDLTSTDYDTAADATESSYFSAQGYLSRE 5246
Db 4878 LKVDKDGSGQVTCVAVNEVGKDSCTAQLNI-----KE 4909
Qy 5247 QEGTSTTDEGOLPOVVEELRDLOVAPGTLAKFOL--KVKGYPAPRLYWKDQOPLTAS 5304
Db 4910 RLIPPSFTK--RLSETVEE-----TEGNSFKLEGRVAGSQBITVAWYKNNIETOPT 4958
Qy 5305 AHIRMTGKILHLEIISVTREDSQYAAIYSNAMGAAYSSARLLAVRGPDEPEEPASDV 5364
Db 4959 SNCEITTFKNNTVLQVRKAGMNDAGLYTCKVSNDSAGSALCTSSIVIK---EPKPPVFDQ 5015
Qy 5365 HEOLVPPMLRFTPKVKYKSSITFSVKVGRPVPTVHMLREBAERGLWIGP--DTPGY 5423
Db 5016 H-----LTPVTVSEGEYVQLSCHVQSGSEPIRIQWLKAGRE-----IKPSDRCSF 5059
Qy 5424 TVASSAQOHSVLVDVGRHQGTITCIASNAAGQALCSASLHVSLPKVBEQEKYKEALI 5483
Db 5060 SFASGTA--VLELRDVAKADSGDYVCKRASNAGSDTTKSKVTIKDKPAPVATPKAAVDG 5117
Qy 5484 STFLQGTQTAISOAGLETASF--ADLGGQRKEEPLA-AKEALGHLSLAEVGTGEEFKULTS 5541
Db 5118 RLFPVSEFQSIRVVEKTTATFIKVG-----DPIPNVKTWKGRQLNQGRVFIHQ--- 5170

Qy 5542 QITEMWSAKITQAKLQVPGGSDSDSKTPS-----ASPRHGRSRPSSSIQ-----ESSSES 5592
Db 5171 -----KGDEAKLEI-----RDTTKTDSGLYRCVAFNEHGEIESNVNLQVDERKKQEK 5217
Qy 5593 BDGDARGEI FDIYVVTADYPLGAEODAITLREGQYVVDLAAHPLRLVTRTKPTKSPS 5652
Db 5218 IEGDLRA-----MLKKTPIKKGAGE-----EEDIMELLKNVDP----- 5253
Qy 5653 RQGVSPAYLDRRLKLSPEWGAAPPEPPGSAVSEYKARLSSVIOBLLSEQAF--VE 5710
Db 5254 -----KEYEKVARMYG--ITDFRGLQAFELLK 5279
Qy 5711 ELQFLOSHLOHLERCHPVPIAVAGQKAVIFRNVDRIGRPHSSFLOELQOCDTDDVAMC 5770
Db 5280 OSQBEETHRL-----IEEIERSEED----- 5300
Qy 5771 FIKQMAPEQVLEFLVGRVQAESVVVSTAIQEFYKVAEALLAGDPSPPPPLQHYLE 5830
Db 5301 -----EKEPEELVSTIQRL-----SQTPE----- 5320
Qy 5831 QPVERVQRYQALLKELIRNKARNQNCALLQOAVVWSALPORAENKLHVSMLMENYPTL 5890
Db 5321 -----VTLIKD-IENQTVLKXNDVAF-----IDIKINYP--- 5349
Qy 5891 BALGEPTRQGHFIWNEGAPGARMPWKGNRHVFLFRNHLVICKPRDRSDTDTVSVPFNM 5950
Db 5350 ----- 5350
Qy 5951 MKLSIDLNDQVEGDDRAFEVWQEREDSVRYKVLQARTAIKSSW-VKEICGIQORLALP 6009
Db 5351 IKLSWYKTEKLEPSDK-FEISIDG-----RHTLAVKNCQLKQDQNYRLVCGPHITASXL 5405
Qy 6010 VMRPDPFEEELADCTAELGETVKLACRVT---GTPKPVISWYKDGKAVQVDPHHILIEDP 6066
Db 5406 TVIEPAWHERHQDVTLEKQGT---CTMTVQFSPNVKSEFRNGRILKPOGRH-KTEVE 5460
Qy 6067 DGSALLILDSITGVDSQYMCFAASAAGNCSTLKLIVQVPP-RFNVKVRASPFVEGEDA 6125
Db 5461 HKVHKLTIADVRAEDQGOYTC---KYEDLETSAELEAEPFOQTKRIONIVSEHOSA 5516
Qy 6126 QFTCTIEGAPYQIRWYKDGALLTGNKFOQLTSLEPRSGLLLVIRAAASKEDLGLVECELV 6185
Db 5517 TFECEV-SFDDAI VTWYKGTTELTSQKYNFRNDGRCH--YMTINHTPDDEGVY--SVI 5571
Qy 6186 NRL---GSARASAEIRIQSPMLQAOEQCHREQLVAADVDTTLERADQVTSVLKRLLAGPK 6242
Db 5572 ARLSRGEARSTAEYL-----TTKEIKLELKP- 5599
Qy 6243 APGSTGDLTGPGP-CPRGAPALQETGSPVPTGTSEAPVPRVPQPLHHEGPPQBEA 6301
Db 5600 ---PDIPDSRVPIPTMPIRAVPPEBI---PPVW---APPVPLLPTP-----EKKKPP- 5643
Qy 6302 IARQEWTPVTRMGAAMPAGAGTGELLDVHSHVVRETTORTYTYQAITHTARPSPMQV 6361
Db 5644 -----PKRIE-----VTKAVKQDAKV-----VAKPEM-- 5668
Qy 6362 TIEDVQAGTGAQFAEIIIEGDPQPSVTWYKDSVOLVDSLRSQQOEGTTSVLVR--HV 6419
Db 5669 -----TPREEIVKPPPTTLIPAKAPEIID---VSSKAEVKIMTIRKKEV 5713
Qy 6420 ASKAGVYTCIAQNTGGQVLCABELLVGGDNEP--DSEKSHRRKLHS-PYEVKEEIGR 6476
Db 5714 QKEEAVY-----EKKQAVHKEKRVFTIESPEEPYDELEVPYTPPEPQYVEEDDYE 5767
Qy 6477 GVFGPVKRVQHKNGKILCAAKFIPLSRSTRQAQVREDDILAALSHPLVTGLLDQETRKT 6536
Db 5768 EIKVEAKKEVEHEWE-----EDPEEQEYIERE-----EGDV----- 5799
Qy 6537 LILILECSSBELDLRLYKGVWVAEYKVYIQOLVEGLHYLHSH-----GVLHLDIKP 6590
Db 5800 -----EGEEWEAAQEREVIQVQKEVYEE-----SHERKVPKVPKPKAPP 5841


```
QY 6591 SNILMVHPAREDIKICDFGFAQNIPTAELQFSQYSGSPFVSPBIIIOQNVPVSASDIWANG 6650
Db 5842 PPKVIKKPVIEKIE-----KTSRMBEEKVQTKV--PE-VSKKIIVPKP----- 5883
QY 6651 VLSYLSLTCSSPAGESDRATLLNVLEGVSVSSPMAHLSDEKDFIKATILQAPQARP 6710
Db 5884 -----SRTVQBE-----VIEVKVPAVHTKKMVISSEKFFASHTEEEVSVTV 5927
QY 6711 SAAQCLSHFWFLKSPAEBAHPINTKQLKFLLRGRWRSLSMSYKILVMSRIPBELLRGP 6770
Db 5928 EV----- 6770
QY 6771 PPSPLGVARHLCRDTGGSSSSSSSSDNELAPFAKSL--PPSPVT----- 6815
Db 5951 PKVPEL-----PEKPAPEVAVPPIPKKVPDPKVPPEVPPKVPPEKKP 5995
QY 6816 -----HSPLLHPRGFLRPSASIPPEAEASERSTEAPAPASPEGAGPAAQCVPRHSVI 6870
Db 5996 VVPVKKEPAAPKVPPEPKVPPEKIPVPVAKKKEAPPKV-----PEVQKGVVTEEKI- 6050
QY 6871 RSLFYHQAGESPEHGA-----LAPGSRHRHARRRHLLKGGYIAGALPGLREPLMEHRV- 6923
Db 6051 -----TIIVTQREESPPAVPEPKVPPEKVPPEKVEEVPVPPKVPALP--KKVPPEKVA 6106
QY 6924 -----LESEAAAREOATLLAKAPSFETALRUPASGTHLAPGH--SHSLEH 6966
Db 6107 VVPVAKKAPPRAEVSVKTVVEEKRFVAEEKLSPAVQFVETVRHRSABEEMSYSSEE 6166
QY 6967 DSPS-----TPRPSSEACGE-----AQLPSAPSGGAPIRDMG 6999
Db 6167 EGVSI SVYRBEEREBEBAEVTYVNMEEPEBYVVEEKLHLSKRVAEBA-----EVT 6220
QY 7000 HPQGSK-----OLPSTGCHPGTAQ--PERPSDPSPMGQ--PAPCHPKQSGAPQEGCSHPA 7052
Db 6221 ERQEKIVLKPAPKIPAKIEPPPAKVPPEAKKIVPEKKVPAPV--PKKSKVPPKVPPEPK 6278
QY 7053 VAPCPGSGFPFGSCK-EAPL-----VPSSPFTLQGOQAPAPAKASPPLDKMG- 7099
Db 6279 --KPVEKKVPKVIKMEEPPLPAKTEKMQITQEEKVLVAVTKKEAPPKARVPPEPKRAV 6337
QY 7100 PGDISLPRKPGPCSSQSGASQASSQVSSLRVGSSQVGTGEPGSLDAEGWTQE--AED 7157
Db 6338 PEEKVLKLPKREE--EPKAVTTEFRKRVVKEEKVIEAPKREPQDIKEVTIMEEKERAYT 6396
QY 7158 LSDSTFTLQRPQEQATMRKFSLGGRGYAGVAGYGTFAFGDAGGMLGQGPWARIANAV 7217
Db 6397 LEEEAHSVQREEEYEEYDYKEFEYEPTTEYDQY----- 6433
QY 7218 SQSEEEQEEAPAESQSEEQQAPAESEPLQVSAAPVPEVGRAP-----TRSSPE---PT 7269
Db 6434 --EYEEEREYERYEEHBEYITEPEKPIPVKVPPEBPVTKPAPKAPKVLKKAAPPEEKVPV 6491
QY 7270 PWEDIGQVSLQIRDLSGDAEAADTISLDISEVDPAYLNLSDLYDKYLPPEFMIFRKVP 7329
Db 6492 P-----IP 6494
QY 7330 KSAQPEPPSMAEEELABFPETWPWGLGPHAGL-----EITEESEDV-----DALLA 7379
Db 6495 KKLKPPPK--VPPEPKVPEEKI-----HISITKREKEQVTEPAKVPMPKRVVA 6544
QY 7380 EAAVGRKKNWSPKSLHFPGOR--HLPLD--EPABLGLRERVKASVEHISRL----- 7429
Db 6545 BEKVPVPRKEVAP-----PVRVPEVPEKELEPEEVAFEVEEY--VTHVEEYLVVEEBEY 6594
QY 7430 -----KGRPEGLEKEGPPRKKP 7446
Db 6595 IHBEETFEETEEVVPVIVPVKVPPEVPRKVPPEKKVPVPPKKKEAPPKVP 6644
```

RESULT 3
T42633
connectin/titin - chicken (fragment)
C:Species: Gallus gallus (chicken)

```
C>Date: 11-Jan-2000 #sequence_revision 11-Jan-2000 #text_change 09-Jul-2004
C/Accession: T42633
R,Yajima, H.; Ohtsuka, H.; Kawamura, Y.; Kume, H.; Murray, T.; Abe, H.; Kimura, S.; Ma
Biochem. Biophys. Res. Commun. 223, 160-164, 1996
A>Title: A 11.5-kb 5'-terminal cDNA sequence of Chicken breast muscle connectin/titin re
A/Reference number: 222221; MUID:96254045; PMID:8660363
A/Accession: T42633
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-4162 <YAJ>
A/Cross-references: UNIPROT:Q98918; EMBL:D83390; NID:gl513029; PIDN:BAAL1908.1; PID:gl51
A/Experimental source: breast muscle
C/Keywords: skeletal muscle

Query Match 6.4%; Score 2629; DB 2; Length 4162;
Best Local Similarity 23.4%; Pred. No. 1.3e-68;
Matches 1110; Conservative 694; Mismatches 1874; Indels 1074; Gaps 172;

QY 9 APRELTRPKAFVSV--GKDATLSQIVGNPTPQVSWKDDQPVTA-----GARFRLAQDGD 63
Db 5 APTF-TQPLQSVVALEGSAAATFEAHISGFVPVEVSWRQGVLSAATLPGVQISFS-DGR 62
QY 64 LYRLTILDLAGDSQYVCRARNAIGEAFAAAGLOVDAEACACAOAPHLLRPTSIHVRE 123
Db 63 A-KLVIPSVTEANSGRVTIQATNGSGQATSTAEALLVTAAGTA-----PPNFSQRLQSMARQ 117
QY 124 GSEATFCRVGSGSPRAVSWSKQGRRLGEPDGPVRVVEELGEASALRIRARPRDGGTYE 183
Db 118 GSQVRLDVRVTGTPPVVKFYRDGVEI--QSSPDFQILQEGDLYSLIAEAYPDSGTYS 175
QY 184 VRAENPLGAASAAALVVDSDAADTASRPGTSTAAHLAHLQRRBA----- 229
Db 176 VNATNVGRATSTAELLIQEEEAVPAKK-TKTIIVSTAQISQTRQARIKKIETHFDARS 234
QY 230 -----MAEGAPA-----SPSPSTGTRCTCTVTEGKHARLSCYVTGEPKPEVWKDQGLV 278
Db 235 LTVSEVMYIEGAAQOLPHKAPPRMPRP-----TSKSTPPPVITAK-AQMA 279
QY 279 TEGRRHVVEDAQENFVKILFCCKQSD---RGLYTCASNL-----VGQTVSSVLVVVRE 330
Db 280 RQQSPSPV-----RQSPSPVHVRAPTPSPVRSVSPAGRISTSPRPVKS 324
QY 331 PAVPFKKR---LQDLEV-----REKESATFLCEVPQP--STEAAWFKEETRLWASAKYGI 380
Db 325 PS-PIRAQVVTPGAELVPPWROEGYATAEAOKMETRVSTSAFEIRTEER-W-EGRYGL 381
QY 381 EEEGTERRLTVRNVSAODDDAVYICETPEGSRTVAELAVQGNLLKRLPKRTAVRVDGTAMF 440
Db 382 QEQ-----VTISGAAA-----GEVAAGAKEVREKEPTPV-----PTVII 416
QY 441 CVELAVPVGPHVHLNRNOEEVVAGGRVAISAEGRHTLTITISQCCLDEVGQVAFMAGDCQTS 500
Db 417 ATDKA-----KQERISTAREEISA---RH-----EQV-----HVS 444
QY 501 TRFCVSAAPRKPPLQPPVDPVVKVARMESVILSWSPPHGERPVTIDGYLVEKKKLGTYTWM 560
Db 445 HEQIEAGKRAEAVATVAAVDAQVRVS-----PMETE-----QVDETYVKKKTL-BYGY 492
QY 561 -----IRCHEAE---WVATPELTVDVAEENFGQFVRVSNALNSFGQSPYLEPPTVHLAPKL 613
Db 493 KEHAVKDHEAOAEHHVAKEVK-----TIVVPEP----- 521
QY 614 AVRTPLKVAQAVEGGEVTFVSDLTIVASAGEWFLDGOALKASSVVEIHCDRTRHTLTIREV 673
Db 522 -----KHIPAAEKKEVHVSTEIK-----RET 542
QY 674 PASLHGALQKFPVANGIESSIRME-----VRAAPGLTANK-----PPAAAAREVLARLHBEAQL 726
Db 543 EAK-----IEKTHIEHPRPRTASPHFTSKIAVPKPDHTYEVSIAGSAMATL 590
QY 727 LAELSDQAAAVTWLKGRTLSPGPKVEVOASAGRRVLLVRDVARDDAGLVECVSRGGRIA 786
Db 591 EKELSATSAQKTKPKVPKPOLKP-HEVKTKP----- 621
```


Db 2284 E--SVKPSAIVTGKPNPNTWYLNKKLQSQSEVVKVYVHETGKTSIRIQKPLMHN--- 2338
Qy 2052 ADRGFFGCETP-----DDKTOA-KLTVEM--ROVRLVRGIQAVEARE-OQT 2093
Db 2339 ---GTRIVEAENVSGKVQATQALKVDKTEVPKFTTNMDRQVK-----EGEDVKFT 2387
Qy 2094 ATMEVOLSHADVGDGWRDGLRPOQGTCHLAVRGMHTLTLSGLRPDSGLMVPKABG- 2152
Db 2388 ANVE---GYPEPSVAMTLNGEPVSKHPNITVTDKGEHTIEISAVTPEQAGELSCATNP 2444
Qy 2153 VHTSARLV-----VTELPVSFRPLQDVVTEKEKVTLECELS--RNVNDRVWLKDGVE 2204
Db 2445 VSGKRDVLAVKKVGDAP--TFAKNLEDRLITTEGELTMDAKLNVKPKPKITWLKDGVE 2503
Qy 2205 LRA-GKTMIAAQAQGRSLIYRCBFADQGVVVCDAHD-----AQSASVKVQGRVTYLIY 2259
Db 2504 ITSDGHYKIVEEDGSLKLSILQTKLEDXGRITIKAESEFGVAECASLGV-----VK 2556
Qy 2260 RVLVADAGEIOFVAENAESRAQLRVKELPVLTVRPLRDKIAMEXHRGVLECOV---SRA 2316
Db 2557 GRPMAPKA---FQSDIA-----PINLT-----EGDTLECKLLITGDP 2590
Qy 2317 SAQVRWFGSQBLQPGPKVELV--SDGLYRKLIISDVHAEDEDTYTC---DAGDVKTSQA 2371
Db 2591 TFFVWYIQTQVLCATEDETEISANGVY--TMKIHGVTADMTGKIKCVAYNKAGEVSTEG- 2648
Qy 2372 PFVEBQSITIVRGLQDVTMEPAPAFWE---CETSI-----PSVRPPK 2411
Db 2649 -----PLKVVAIPVEFETSLCDATCREGDTLKLRAVLLGEPEPV--S 2690
Qy 2412 WLLGKTVLOAGNVGLEBGTVHRLMLRRTCSMTGPHVFTV-----GKSRSAKRLVSDI 2467
Db 2691 WTVNGKLEESNIKIHSEKGTYYTIIKDITCDYSGVVCEAINEYGHATSEATL----- 2745
Qy 2468 PVVLRPLREP-----KTG-RELQSVVLSCDFRPAPKAVOMYKDDTPLSPSEKFK 2515
Db 2746 -LVLPRGPPDPLEWLSNVARTGTVKVVVFTGDPKP---SLTWYNNKEILNSDLYT 2801
Qy 2516 MSLEQMAELRILRMP-----ADAGVYRQOA-----GSAHSSTEVTVARE 2557
Db 2802 IVTDKKTSTLTINSFPDVHVGEEIICAENDAGEVSCVTANMITYTSDMFSESESAQAEE 2861
Qy 2558 -----VTVTCP-----LVLKSIORADAG-----IVRASSLKV 2634
Db 2862 FVGDDLTDESLREEMHRTPTVMAKPFITKIKDTKAKKHSVAFECVY--PDTKGVCCK 2919
Qy 2590 WSLANGPLYNDSFHSHKGRHT-----LVLKSIORADAG-----IVRASSLKV 2634
Db 2920 WLKDGKEI-----ELIARIIVQRTGPEGHITQELVLDNVTPEDAGKYTCIVENTAGKD 2973
Qy 2635 STSARLEV-----RVKPVVFLKALDDLGAERGTALAOCEV--SDPEAHVVMRKDGV- 2684
Db 2974 TCEATLTVIESLEKSEKKAPEFIVALQDKTKTSEKVVLECKVIGEBPKPKVSMHLDNVS 3033
Qy 2685 -QLGPSDKVDFLHTA-----GTRGLVVHDVSPEDAGLYTCHV-----GSSETARVRVH 2732
Db 3034 REKNPSEKTIQESITVESVEGVERVTITSELSHOGKYTCIAENTEGTSKTAFLTVQ 3093
Qy 2733 DLHVGITKRLKTMVELEGESCFEVLGSHESASDPAMMTVGKGT-----VGSSSRF--- 2783
Db 3094 GEAPVFTKELQNKELSIGELVLSCSVKSGSQPHVDYFSPSETKYVETKITSSRIAIEH 3153
Qy 2784 -QATRQGRKYI-LVVREAPSDAGEVFSVRLGTSKASLIVRERPAALIKPLEDOWVAPG 2841
Db 3154 DQTNHMRWVISQITKEDIVSYKAIATNSIGTATSTSKITTKVEAPVFEQGLKKTYSVKEK 3213
Qy 2842 EDVELRCFLSRAGTPVHMLKDKRAIRKSQYDVV--CEGTWAMLVIRGASLKDAGEYTC 2899
Db 3214 BEIKHEVKGGSAPDVEWFKDDKPVSEDNHEMKKNPETGVFTLLVVKQAAATTDAGKYTAK 3273
Qy 2900 VEASKSTASLHVEKAN-----CFTTEL--TNLQVEBKGTAVFTCKTEHPAAVTTWRKG 2951
Db 3274 ASNPAGTAESAESAEBVTQLEKPTFVRVRELVTTEVKINETATLTVTVKGV-PDPSVEMKLD 3332

Qy 2952 LLELRASGHQPSQ-EG-LTLRLTISALEKADSDTYTC-----DIQQAQSRALLVQGRRV 3005
Db 3333 GQPQTDSSHVIKVEGSGSYITIKOARLEDSKYACRAINPAGEAKTEANFAVKNLV 3392
Qy 3006 --HIIELEDVDDVOEGSSATPRCISIPANYEPVWFIDKTPHLHANELNEIDAQPG-GVHV 3062
Db 3393 PPEFVEKLSPLVEKEKSTLTSVKVVGTPPEPSWFKDDTPISIDNVHVIQKQTAVSFS 3452
Qy 3063 LTLQALOKSGTIYF-----EAGDQASAAALVTEK--PSVFSRELTDAITIGEDITL 3115
Db 3453 LTIINDAQDVG-LYSCRAARNEAGEALTANFGIIRDSIPPEFTQKURPLEVREOETL 3511
Qy 3116 VCETSTCIDP-MCWTKDQKTLRGSARCQLSH-----EGHRAQLLITGATLQDSGRYKC 3167
Db 3512 KVTYIGTVPNVNFWFKDKPI---NIDNSHIFAKDESGHHT-LTIKQARGEDVGVTYTC 3566
Qy 3168 ----EAGACSSSIVRVHAR---PVRFOEALKOLEVLEGGAAATLRCVLSVAAP-VKWY 3219
Db 3567 KATNEAGEAKTTANNAVQEEIEAPL-FVQGLKPYEVEQGKPAELVVRVEGKPEPEVKWFK 3625
Qy 3220 GNNVLRPGDKYSLRQEG-AMLELVVRNLRPQDSGRYSC---SFGDQTTSATL----- 3267
Db 3626 DGVPIADNQHVIIEKKGENSHLTVIKDTNNADFGKTCQATNKAGKDETVGEUKIKPYS 3685
Qy 3268 ---TVTALPAQFTIGKLRNKEATEGATATLRCELSKTA--PVEWRKGSSETLRDGDYCLR 3321
Db 3686 FEKQTAEEVPLFTIEPLKETFAVEGDTVLECKYKNKESHQIKPFKNDQPVIEIQHMQLB 3745
Qy 3322 --QDGMCELQIRGLAMWDAAYSC---VCGEBRTSASLTI----- 3357
Db 3746 VLEDGNI-KLTIONAKKEDVGAYRCEAVNVAGKANTNADLKIQPAAKVEEHVTDSEGOLE 3804
Qy 3358 -----RPMPAHFIIGRLRHQBSIEGATATLRCELSKAAP---VSWRK 3395
Db 3805 EIGQFETVGTASKTDTGKAP-EFVELLRSCVTVEKQAILKCKV-KGEPRPKIIMWTK 3862
Qy 3396 -GRESLRDGRHSLRQDGVCELIQICGLAVADAGEYSCVCGEERTSA-----TLTKAL 3448
Db 3863 EGKEVEMSAKVRABHKDDGTLTLFDNVQTADAGEYRCEAENEYGSATWETPIIVTLEGA 3922
Qy 3449 P-----AKTEGLRNEEAVEGATMLWCELS-KVAP-VWEWRKGPENLRDGDYCLR--Q 3498
Db 3923 PKIDGEAPDLQPVKPAVTVGTAVLEKIGISKPKPSVKWYKNGEELKPSDRVKIENLD 3982
Qy 3499 EGTRCELQICGLAMADAGEYLCVCGOE---RTSATLITRALPAR----FIEDVKNOEAR 3550
Db 3983 DGTQ-RLTVTNAKLDDMDEYRCEASNEFGDVWSDVTLTVKE-PAQVAPGPFKELSAIOVK 4040
Qy 3551 EGATAVLQCELSAAP-VWEWRKGSSETLRDGR--YSLRQDGTKCELOIRGLAMADTGEY- 3606
Db 4041 ETETAKPECKVSGTKPDVKWFKDGTPLKEDKRVHFESTDDGTQ-RLVIEDSKTDDQNYR 4099
Qy 3607 ---SCVCQGBRTSAMLTVRALP---IKFTEGLRNEEAEGATAVLRCEL-SKMAPVEMWK 3659
Db 4100 IEVSNDAGVANSKVPLTV--VPSETLKIKGLTDVNVYTOGTKILLSVEVEGKPKTVKWK 4157
Qy 3660 GHETLRDGRHSLRQ-DGARCELQIRGLVAEDAGEYLCMCGKE----RTSAMLTVR--- 3710
Db 4158 GTETVTSOQTKIVQVTESEYKLEIESAEMSDTGAYRVVLSTDSFSVESATVTVTKAAE 4217
Qy 3711 --AMPSKFIIEGLRNEEAEGDTATLMCEL-SKAPVWEWRKGHETLRDGRHSLRQDGSRC 3767
Db 4218 KISLPS-FKKGLADQSVPGKTPLVLEVEIEGKPKDVWKYKNGEIKDQKVEDL--GNGKY 4274
Qy 3768 ELQIRGLAVVDAGEYSCV---CGOERTSATLTVRALPARFIEDVKNOEARREGATAVLQC 3823
Db 4275 RLTIPOFQEKDQVGEYSVTAAANEAGEIESKAKVNVSAKP-EIVSGLVPTTQGETATFNV 4333
Qy 3824 EL-SKAAPVWEWRKGSSETL-----RGGDRYSLRQDQTRCELQIHGLSVADTGEYSCV-- 3873
Db 4334 KVGKPVKVKWYKNGKKEIPDAKTKDNGDGSYS-----LEIPNAQVEDAADYKVVS 4384

Db 6038 -----EVNGEPEE-----CKPR----- 6049
QY 5945 YVFRNMKLSIDLNDQVEGDDRAPEVQWQEREDSVKRLQALQARTALKSSWKEICGQQ 6004
Db 6050 -TRGLYNS-----IHE----- 6061
QY 6005 RLALPVMRPDPFEELADCTABELGETVKLACRVGTGPKPVIWYKDGKAV---QVDPHHI 6061
Db 6062 -----GNVEMIVCATGTPPTVWXYKDGQEIIVGDPGDKRV 6098
QY 6062 LIEDPDGSCALLDSLTGVDSCQYMCFAASAAGNCSITGLKILV-----QV 6106
Db 6099 IFTDERGIHLLVIVNASPDDEGEYSLEATNKLGSATGSLNIIRPRIADADERGMPF 6158
QY 6107 PRFRVNMKVRASPFVEGEDAQFTCTTEGAPYPOIRWYKDGALLTGNKFTILSEPRSLLV 6166
Db 6159 PGFVRQLKNKHFVNHMPTIFDCLVVGHPAPEVFNHNGKLVPGRIK-IQSCGGGSHA 6217
QY 6167 LVIRAASKEDLGLYBCELVNRLGSARASAEIRIQSPMLQAOBQCHREOLVAAVEDTTLER 6226
Db 6218 LIILDTTLEDAGEYVATAKNHSGSASSAVLDTVPFL-----DSIKFN 6261
QY 6227 ADOEYTSVLKRLG-----PKAPGSTGLTGP-----PCPRGAP----- 6262
Db 6262 GEIDVTPYLTEYGFKKLNTASLPTPPDRGPFKEVTHYLTSLWIPTKRAPPYQVTV 6321
QY 6263 -----ALOETQSOPPV-----GTSE-APAVPP----- 6284
Db 6322 VIEIRELPEKQWLSLEYNIPEPVCKVRNLELGSQFVRRAENIYIGIDPSPASPRLM 6381
QY 6285 RVPQPLLHE-----GPEQEPBAIARAQE-----W-----TYQAIDHTA----- 6354
Db 6382 APPQPVFORRTNKVPLLDPAEAKALDMRYSOYACAPWFSPGVVEKRYCAENDTLTIVL 6441
QY 6313 RMEG-----AAMPAGTCELLWDVSHVURETQT-----W----- 6354
Db 6442 NVSGPDPDIKWKFRG-----WDI-----DTSSEPTSKCVYTYGSETTLAIGTFSKEN 6490
QY 6355 -----RPPSMQVETIEDVQAQGTGAQ-----FEAITEGDPQ 6386
Db 6491 VQYQCFKANDYGAQQIMVDLATRPNFIQLVN-----KTFSSAQPMRMDVRVDGEFP 6546
QY 6387 SVTWYKDSVLVDSRSLQQOEGT-TYSLVLRHVASKDAGVYTCIAQNTGGQVLKABELL 6445
Db 6547 ELKMKMKWRPIVESSRIKFPVQDGYLCSLIINDPWRDSGIYSCVAVNDAGATTSCTVT 6606
QY 6446 VLGGDNEPDSKQSHR-----RKLHSFYEVKEE 6473
Db 6607 VEAEGYNDVLPFRRTVETIESRRVRELYEISEK 6639

RESULT 5
T43290
hemictentin precursor - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 11-Jan-2000 #sequence revision 11-Jan-2000 #text_change 09-Jul-2004
C:Accession: T43290; T20993; T24734
R:Voegel, B.E.; Hedgecock, E.M.
submitted to the EMBL Data Library, June 1998
A:Description: Hemictentin is required for hemidesmosome mediated cell adhesion and germ-
A:Reference number: 222396
A:Accession: T43290
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-5198 <VOG>
A:Cross-references: UNIPROT:O76518; EMBL:AF074901; PIDN:AAC26792.1
R:Suiston, J.
submitted to the EMBL Data Library, December 1994
A:Reference number: 219355
A:Accession: T20993
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5198 <WIL>

A:Cross-references: EMBL:Z47068; PIDN:CAA87336.1; GSPDB:GN00028; CESP:F15G9.4b
A:Experimental source: clone F15G9
R:Kershaw, J.
submitted to the EMBL Data Library, December 1994
A:Reference number: Z19929
A:Accession: T24734
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-5198 <W12>
A:Cross-references: EMBL:Z47070; PIDN:CAA87345.1; GSPDB:GN00028; CESP:F15G9.4b
A:Experimental source: clone T09B9
C:Genetics:
A:Gene: him-4; F15G9.4b
A:Map position: X
A:Map position: 85/1; 120/1; 334/3; 370/1; 477/2; 606/3; 664/1; 935/3; 977/1; 1051/3; 1184/3;
; 2512/3; 2593/3; 2699/3; 2759/1; 2852/1; 2899/3; 2913/3; 2941/1; 2967/3; 2991/3; 3033/1;
1; 4225/1; 4361/1; 4408/1; 4456/1; 4498/1; 4647/3; 4836/1; 4902/1; 4964/1; 5034/1; 5100/

Query Match 5.0%; Score 2048; DB 2; Length 5198;
Best Local Similarity 21.4%; Pred. No. 1.1e-51;
Matches 1187; Conservative 773; Mismatches 2284; Indels 1292; Gaps 254;

QY 286 VYEDAQ-----ENFVLKILFKQS-----DRG-----LYTCTASNLVGQTY-- 321
Db 136 VFTDARSKDHLDEVLTNIQEKQSSVYFVMTGDCGNRTHPGFRTYKIAAASFGVPHL 195
QY 322 -----SSVLVVVREPA-----VPPFKRLQDLEVR-----EKES 349
Db 196 EKSDVSVLEVYRVHVAQKKVHLMYEAERGTVSRNIPVDKHLSELTISLGGDKSDN 255
QY 350 ATFLCEVPQPSAEAFKEETRLWASAKYGIIEEGTERRLTVRNVSDADDVAVICETPEG 409
Db 256 LDIVLRDPEGRT-----VDKRLYS-----KEGGTIDLKNVLIKRLKDPSPGVMTVNTN 303
QY 410 SRTVAELAVQGNLLRKLPRKTAVRVGDTCFVELAVPGVPHVLRNQEVEVAGRVAS 469
Db 304 SRLKHTIRVGHGAVDPKYGFASRPLDR-----IELAR-PPV---LNQDTYLLINMTGLI 355
QY 470 AEGTRHRLTISQCCLEDDVGQVAFMAGDCQTSFRCVSAAPRKPPLQPPVDPVVKARMESSV 529
Db 356 PPGT-----VGEIDLVDYHGHSLYKAVASPHRTNFMNYFAGFPV----- 394
QY 530 ILSWSPPHGERPVTIDGYLVEKKKGLCTYTWIRCHEAEWATPELTADVADAEQNFQFRV 589
Db 395 -----PPKGLGFVRVQGY-----DEDNYSFMR 416
QY 590 SALNSFCQSPYLEPFGTVHLPAPKLAVRTPILKAVQAVEGGEVTFVSVDLTVASAGE--WFLD 647
Db 417 IAPTAI-----GSIIVGGPRAFMSPIH---QEFVGRDLNLSCTVESASAYTIYVWKT 465
QY 648 GOALKASSVYEHCDRTRHRLTTRVFPASLHGA---OLKFFVANGIESSI--RMEVRAAPG 702
Db 466 GEDIIGGLFYHNTDTS--VWTIPEL--SLKDAGEYECRVISNNGNYSVKTRVETRESP- 520
QY 703 LTANKPPAAAREVRLARLHEEAQLAELSDQAAA---VTWLKDGRTISLPQPKYEVQASAG 759
Db 521 -----PEIFGVRNVSVPLGEAFL--HCSTRSAGEVIRWTRYGATVFNQNTERNPTNG 573
QY 760 RRVLLVRDVARDDAGLYECYSR--GGRIAYQLSVQGL-----ARFLHMDMAGSCVDVAVAGG 813
Db 574 --TLKIHVTRADAGVYECVARNAGGMSRKMRLDIMEPPSVKVTPODVVFNMRGV--- 628
QY 814 PAQPECET-SEAHVHVHVKDGMELGHSGERFLQEDVGTTRHRLVAAATVTRQDEQTYSCRV 872
Db 629 --NLSCEAMGDPKEVHVHVKFGRHLLNDYKVVQGD---SKFLYIRDATHHDEGTYECRA 683
QY 873 -----GEDSVDFRLRVSEPKVVFVFAKEQLARRKLOAPAGASATLSCEVAQAQ--TEVTWY 924
Db 684 MSQAGQARDTTDLML-ATPPKV-----EIIQNMVWGRGDRVSVFECKTIRGKPKPKIRWF 737
QY 925 KDGKLLSSSSKVCMEATGCTTRRLVVOQAGQADAGEYSC-----BAGGQRLSFHLDV-KEPK 979
Db 738 KNGKOLIKPDYIKINEG---QLHIMGAKDEDAGAYSCVGENMAGKDVQVAVNLVSGRVPT 794

1618 HVAEPKVVFAKEQ--PAHREVQAEAGASATLSCVAQAQTEVTVYKDGKLLSS--SSKVR 1673
1400 -----IVDAKNKFVIALKDTVEIEKDDVTLTCQTKDTKPGIWFNGKQISSMPGGKPE 1453
1674 VEAVGCTRLRVVQQAQADAGYSCEAGGQRLSFRHLHVAELE--PQISERPCR-----RE 1726
1454 TQSRNGTHTKIGKIEEMNEADYVEIDQAGLRSCNVTVLEAEKRPILNWKPKKIEAKAGE 1513
1727 PLUVKEHBDIILTATLATPSAATVTLKDGVBIRSKKH--ETASQGDTHTLTVHGAQVL 1784
1514 PCVVVPFQIKGTRR-GDPKAQI-----LKNKGPIDEBMRKLVIIKDDVAEIVFKNPQIA 1569
1785 DSAIYSCRVG-----AEGODF-----PVQVEEV 1807
1570 DTGKWALELGNAGTALAPFELFVKDKPKPKPLETKNVTAEGLDLVWGTDPDDEGAPV 1629
1808 AAKFRLLE-----PVCGLGCTVTLACELSPACAEVVMR-----CNTQPRVGRKRPQ 1855
1630 KAYIIEQGRSGNWKVGETKGTDFKVKDLKEH-GEYKFRVKALNECGLSDPLTGE--S 1686
1856 MVAEPVRSLTVLGRADAGYVCESDRDHTSAQLTVSVPRVKFMGLSTVVAEBEGE 1915
1687 VLAKNP-----YGVPGPKPKMDAIDVDKDHCT-----LAMEPPEEDGG 1724
1916 ATFQCVWSPSDVAVVWFRDGLALQPSK-----PAISQSGASHSLTISDLVLEADAGQI-T 1969
1725 API-----TGYIIERREKSEKMDHQVQGTQKPCCBLTDKKKVVEDKEYLYR 1769
1970 VEAEGASSAALRVREAPVLFK-KKLEPO-----TVEBRSSVTELEVLT-RPWPEL 2018
1770 KVAVNKGKPCDPCDHGKPKKAKKASPEFTGGIKDLRLKVGETIKVDVPISGEPLPEC 1829
2019 RWRNATALAPQKNVEIHAEGARHLVLHNVGFADRGFF-----GCETPPDDKTQAK 2069
1830 LMVNVGKPLKAVGRVKMSSERKHKIMKIENAVRADSGKFTIILKNSSGSCD-----STAT 1884
2070 LTVENRQVRLVRLQAVEAREOGTATME-----VOLSHADVCGSWTRDG 2113
1885 VTVVGRPTPKGPDIAVDCADG-ATLSMNPDDDDGDLPTGYIVEAQMDNKGKYEIVG 1943
2114 LRFQOGPTCHLAVRGMHTLTLGLRPEDSGLMVFKAEGVHTSARLVTVELPVFSRPLQ 2173
1944 -----KVDPTNTTLKNGLR--NKGNYKFRVAVNNEGE-----SEPLS 1980
2174 -DVWTEKEKVTLECELSRP-----NVDVRW--LKDQ-----VELRAGKTMAI 2213
1981 ADQYTOIKDPWDEPKGPRPEITDFADRIDIAEPHPKDGAGPIEYIVEVRDPDK-- 2038
2214 AAQACRSLTIVRCFADQGVVVCDAHDAQSSASVKVQGRVTLIYRVRVLAEDAGEIQPV 2273
2039 -----EWKE-----VKRPVPTNAGISGLKEGKEYQFRVRAVNKAGPGQ----- 2076
2274 AENAESRAQL-RVKELPVTVLVRPLDKIAWEKXRGVLEQCVSRASA-----Q 2319
2077 -PSESEKOLAKPKFIAPWL-----KHDNLKSTVAKAGATVRWEVKIGGEPIPE 2124
2320 VRWFGSOELOPKPYELVSD-----GLYRKLIIISDVHAEDE----- 2356
2125 VKWFKGNQLENG--IQLTIDFRKNEHTILCIPSAWRSVDGEYR-LTVKNSHGAEDEKAN 2181
2357 ---DITYCDAGDVTKTSQAFVBEOSITIVRGLQDVTVNMP-----APAWFEC 2400
2182 LTVLRPSKPNGLFVSDVF-----EDNLNLSWPPDDGGEPIEYVEVEKLDATATGRWVPC 2238
2401 -----ETSIPEV-- 2407
2239 AKVKDTKAHIDGLKGGQTVQFRVAVNKEGASDALSTDKDTKAKNPNPEPKGTGPDVVD 2298
2408 -----RPPKWLKGLKTVLQ-----AGNVGLE-----OEGTVH 2434
2299 WDADRVSLEWEPKSGDGAIPITQYVIEKKHGRDQWQECGKVGSGDQTNAEIILGLKEGBEY 2358

2435 RLMLRRTCTMTGTPVHFTVGKSRSSARLVVSDIPVVLTRPLEPKTGRELQSVVL-----S 2489
2359 QFRVAV--NKAGP-----GEASDPSRKVAK-----PRNLKPWIDREAMKITITIKVND 2406
2490 CDF-----RPAPKAVOMYKDDTPLSPSEKFKMSLEGQMAELRILRLMPADPAGVYRCQA- 2542
2407 VEFDPVVRGEPPPKK-EMIFNEKPYD-DQKIRIESEDYKTRFVLRGATRKHAGLTLTAT 2464
2543 ---GSAHSTEVTVPEAREVTVTGPIQDABATEEGHNASCELSHEDEEVEWSL-----NGM 2595
2465 NASGDRKHSVEVILGKPSPLPLEVSNVYB-----DRADLEWKPVEDDGA 2512
2596 PLYNDSPEHSHK-----GRHHTLVLSIQRADAGIVRASLSKYSTSARLEVRVK 2645
2513 FI---DHEIEKMDLATGRWPCGRSET-----TKTVPNLQPHGEYKFRVR 2556
2646 PVVFLKALDDLSAE-----ERGTALQCEVSDPEAHVVMRKDGVLQ- 2687
2557 AVNKEGESDPLTNTAILAKNPYEPVGVKDPKEL--VMDKDHVDLAWNAPDDGGAPIEAF 2615
2688 -----PSD-----KYDFLHTAGTRGLVHVDVSPEDAGLYT 2717
2616 VIEKDKNGRWEALVWPQDKTATVPLNKEGEYQFRISARNKAGTGDPSDPSD----- 2670
2718 CHVGSSEETRARVRVHDLHVGITKRLKTMEVLEGESCFEVLSHESASDPANMTVGGKTV 2777
2671 -RVVAKPRNLAPRIH-----REDLSDTTVKVGATLKFIVHIDGEPADP-VTWSFNGKI 2722
2778 GSS-----SRF-----QATROGRKYILVREAAAPSADAGEVVSFVRGLTSK----- 2817
2723 GESKAQIENEPVSRFALPKALRQSGKYTITATNINGDSTVITINIKVSKTKPKGPIE 2782
2818 ASLIVRERPAIILPLEDO-----WV----- 2838
2783 VTDVEDRATLDWKPPEDDGGEPFIEFYEIKMNTKDGIVWPCGRSGDTHFTVDSLNGDGH 2842
2839 -----APGEDVELRCEL-----SRAGTP-----VHW--LKORK 2864
2843 YKFRVAVNSGSPDPLETETDILAKNPFDRPRPEPTDWDSDHVDLKWDPPLSDGG 2902
2865 A-----IRKSOKYD-----VVCETMAMLVIRGASLKDAGEYTCBV-----E 2901
2903 APIEYQIEKTKYGRWEPAITVPGQTATV--PDLTPNEEYEFVRVAVNKGSPDPSD 2960
2902 ASKTASLHVSEKANCFTTEELTNQVBEKGTAVFTCKTE-HPAATVTVR-KGLELLEABG 2959
2961 ASKAVIAKPRNLKPHIDRALKNLTIKAGQSISFSDVPVSGEPAPTIVTWHWPNREIRNGG 3020
2960 KHQPSQEGTLRLTISALEKADSDTYTCDIGQAQSRALLVQGRVHIIE-----D 3010
3021 RVKLDNPEYQSKLVYKQMERGDSGFTTAKVANNAGEDEATV---KINVIDKTPSPNGPLD 3077
3011 LEDVD-----VQEGSSAT----- 3023
3078 VSDVGHGHTLVNWRAPDDDDGGPIENYVIEKYDVTASGRVWPAKVAGDKTTAVVDGLIFG 3137
3024 ---FRCRISPANYE---PVHWF---LDKTPPL-----HANELNEID---AQF 3057
3138 HEYKFRVAVNAEGESDPLETFTGTTAKDPFKPKGTNAPEITDWDKOHVDLEWKKPAND 3197
3058 GG-----YHVLTLRQLA-----LKDSGTIYF-----EAG-----D 3082
3198 GGAPIEEYVEMKDBFSPFNDVAHVAPAGQTNATVGNLKEGSKYEFIRAKNKAAGLDPS 3257
3083 QRASAAARVTEKPSVFSR-ELTDATITBEGEDTLVCETSTCIP-MCMTKDGKTALRGASR 3140
3258 DSASAVAKARVPPVIDRNSQEIYKVKAGQDFSLNIPVSGEPTPIITWTFTGTPVESDDR 3317
3141 COLSHEGHRAQLLIITGATLQDSGRY--KCEAGGACSSSIVRVA-----RPRVFORA 3190
3318 MKLNNEDGKTFHVKRALRSDGTGTIYIKAENENGTDTAEVKTVVLDHPSPRGLDVTNI 3377
3191 LKOLEVL-----EGGAATLRCLVSSV-AAPVKV-----CYGNN----- 3222

Db 3378 VKQCDLAWKEPDDGGAIEISHYVIEKQDAATGRWTACGESKDTNFHVDDLTQGEHYKFR 3437
QY 3223 ---VLRPGDKYSLRQEGAML-----ELV-----VRNLRQDQSGRYSCS 3257
Db 3438 KVAVNRHGDGDPLEAREAIIAKOPDPRADKPGTPEIVDWDKHDKLTWTPPADGGABIE 3497
QY 3258 FGDQTTSATLTVALPAQFIGKLRNKEATEGATATLRCELSKTAPEVWRKGSSETLRDGR 3317
Db 3498 -GYLVEMETPSGDWVPATVG-----AGE-LTAIV-----DGLKPGQT 3533
QY 3318 YCLRODGAACELQIRGLAWDAEYSCVCGBERTSASLTIRMPAHFTRLRH-----OE 3372
Db 3534 Y-----QFRVKALNAGE-----STPSDPSRTWAKPRLHAPKINRDMFVAQR 3576
QY 3373 SIEGATAT--LRCELSKAAPVEM-RKGRESLRDGRHSRLRDGAACELOICGLAVADAGE 3429
Db 3577 VKAGQTLNFDVNVVEGEPAPKIEWFUNGSPLSGGNTHIDNNNDNNKLTUXTSTAKRDSGK 3636
QY 3430 YSCVC-----GBERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLWCELSKVAPVWRKGP 3485
Db 3637 YKIVATNESGKDEHEVDVNIIDIPGAPGRLRHKDIITKESVVLKWE-----PLDDGGSP 3691
QY 3486 -----ENLRDGRVILRQEGTRCBLQICGLAMADAGEY---LCVCGQERTSATLTI-RA 3535
Db 3692 ITNYVVEKQEDGGRWVPCGETSDTSLKVNKL--SEGHEYKFRVKAVERNQGTSAPLTSDHA 3749
QY 3536 LPARFIEDVKGQEARREGATAVLQCELSAAAPVWRKGSSETL-----RDG---DRYSL- 3584
Db 3750 IVA-----KNPDEPDAP-----TDVTPVDWDKHVDLEWKPDPNDGGAPIDAYIVE 3796
QY 3585 -----RODG--TKC-----ELQIRGLAMADTGEYS-----CVCGQER 3614
Db 3797 KQKQFGDWECARVDGKTTKATADNLTPGETVQFRVKAVNKAGPKPSDPTGNVVAKPRR 3856
QY 3615 TSAMLTVRA---LPKFTTEGLRNEEATGATAVLRCELSKAPVWVWKGHE-TLRDGRH 3670
Db 3857 MAPKLNLAGLLDLRIKAGTPKLDIAFEG-----EPAPVAKWKANDATIDTGARA 3906
QY 3671 SLRDGARCELOIRGLVAEDAGEYLCMC-----GKERTSAMLTVRAMPKFTIEGLRNEEA- 3725
Db 3907 DVTNPTSSAIIHFAVRGDTGVYKIIIVENEHGKDTAQCNVTLVDVPGTPEGPLKIDIEH 3966
QY 3726 TEGDTATLWCELSKAAPVWRKGHETLR-----DGRHSLRQDGRSCELOIRGLAVDA 3779
Db 3967 KEG-----CTLNWKPPTD-NGGTDVLHVIVEKMDTSRGTWQEVGTFPDUCTAKVNKLVP 4019
QY 3780 GEYSCVCGQERTSA-TLTVRALPARFIE-----DVK-N 3810
Db 4020 KEYAF-----RVKAVNLQGESKPLEAEBFIIAKNQFDVPDPVDPKPEVTDWDKDRIDIKWN 4074
QY 3811 QEAREGATAVLQ--CELSKAAPVWRKGSSETLRGGDRYSL--RODGTRECELOIHGLSVAD 3866
Db 4075 PTANNGGAPVTGYIYEKKEGSAIWEAGT--PGTTFSADNLKPGVEYEFVRVIAVNAAG 4132
QY 3867 TGEYSCVCGQERTSATLTVRAPQVFRBPLQSLQAEESTATLQCE-LSEPTATVWVSKG 3925
Db 4133 PSDPSDPTDQITKA---RYLKPKILTASRKIKIKAGFTHNLEVDVDFIGAPDPTATMTVG 4188
QY 3926 GLOLOANGREPRRL-----QGCTAELVLQDLQREDGEYTCCTCGSATSATLTVTAAPVRF 3981
Db 4189 ----DSGAALAPELLVADAKSSTTSIFFPSAKRADSGNY-----KLKV 4226
QY 3982 LRELQHE-----VDEGTAHLCCELSRAGASVEWRKGSLOLPFCYQWY-----QD 4029
Db 4227 KNELEDIAI FEVI VQDRPSA-----PEGPLEVSDVTKDSCVNLWKPYPK 4271
QY 4030 GAAABELVRGVEQEDAGDYTCDTGHTQSWASISVVRPKFKTRLOSLEQETGDIARLCC 4089
Db 4272 DGGABISNVVBEKRTKNTW-----VPVSAFV----- 4300
QY 4090 QLSDAESGANVQWLKEGVELHAGPKYEMRSQCATRELLIHQLEAKDTEYACVTGGQKTA 4149

Db 4301 -----GTSITVPKUTEQHE-----YEFR-----VMAENTFGRSDSLNTDEPVLAKDPFGTP 4346
QY 4150 ASLRVTPEVTVIRGLVDAEAVTADEDVEFS-----CEVSRAGATGVQWC-L 4194
Db 4347 G--KPGRPEI-----VD---TDNDHIDIKWDPDRNGGSPVDHYDIERDKAKTRWIKV 4395
QY 4195 QGLPLQNEVTEVAVROGRIHTLRKGVTPPEDAGTVFHLGNHASSAQLTVRAEVTITLE 4254
Db 4396 NTSVPQGTAFSDTRVQKG--HTYEYRVVAVNKAAGPGOPSDSSAAATAKPMEHAKPFIDL 4453
QY 4255 PLQDVQLSEGODASFQCELSRASQOEARWAL-GCVPLQANEMNDITVBEQGLTLLHLTLHKV 4313
Db 4454 DGKEFRVKAGPLVITIPFTASPOPDISWTKEGKPLAGVETTDSQTK-----LVIPST 4507
QY 4314 TLEDAGTVSFHVGT--CSSEAAQLKVTAKTNTVVRGLENVLEALGEGEALFECOLSOPVAAH 4371
Db 4508 RRSDSGPVKIKAVNPYGEAEANIKITV-----IDKGAPE-----NITYPAVSRH 4552
QY 4372 TWLDDDBPVRTSENAEVVFPFENGRLHLLLLKNLRPQDSRVTFFLAGDMVTSAPLTVRGWR 4431
Db 4553 TCTLNWDAPKDDGGAETAGYKIEYQEV-----GSOIWDKVPGLISGTAYTVRG-- 4600
QY 4432 LEILEPLKNAAVRAGAQAARFTCTLSEAVPVGCEASWYINGAAV---QPDSDSWTMTVTDGSH 4488
Db 4601 -----LEHGQQIRFRIRAEANVGLSD---YCGVFPVVIKDPDP-- 4636
QY 4489 QALLLSAQPHHAGEVTFACRDAVASARLTVLGLPDPPEDAEVVAHSHSTVTLSWAAPMS 4548
Db 4637 -----PGAPSTPEITGYDTNQVSLAWNPPRD 4662
QY 4549 DGGGLGCGYRVE--VKEGATGQWR-LCHELVPGBECVVDGLAPGETTRFRVAAGVPVAG 4605
Db 4663 DGGSPILGYVVERPEKRG--GDWAPVPMVVKTECIVPGLHENETVQFRVAVNAAGHG 4721
QY 4606 EPVHLPTQVTLAEPPKVPPOPSPESROVAAGEDVSLLEBVAAEAGVIVHKGMRIOQ 4665
Db 4722 EPSNGSEFVTC-----RPYVEKPGADAPRG-----KITKNSAELTNWR---PLRD 4765
QY 4666 GG-----RFEVSGROOQLVVIKGFTAEDQGEYHCG 4696
Db 4766 GGAPIDGYIVEKKLGDNDWTRCNDKPVDRDTAFEVKNLGEKE-----EYEF 4812
QY 4697 LAQOSICPAATFOVALSPASVDEAPQPSLPPEAAQEGDLHLLWEALARKMRMSREPLD 4756
Db 4813 V-----IAVNSAGEGEPKSPS-----DLVLIIEQPGR-----PIFD 4843
QY 4757 --SISELPEEDGRS--ORLPQEAEEVAP--DLSEGYSTADELART-----GDADLSHTS 4804
Db 4844 INNUKDTVRAGETIQIRIPYAGGNPKPIIDLFGNSPIFENERTVVDVNFGEIVITTG 4903
QY 4805 SDESRAGTPTSLVTLKKAGR-----PG--TSPL-ASKVGAPAAP-SVKPOQ 4848
Db 4904 S-KESDAG-PYKISATNKYKDKTCLNVFVLDAPGKPTGPIRATDIQADAMTSLWRPPKD 4961
QY 4849 Q-----EPLAARPPGLDSTKDLGDPMDKAAVKQAAFKGYKVRKEMQOEGPMESH 4902
Db 4962 NGSDAITNYVVEKRTPGGDWVT--VGHPVGTTLVRNLDANTPYEFRVRAENQYQ----- 5014
QY 4903 TFGDTEAQVGDALRL-----ECVVASKADVRRARWLK---DG----- 4935
Db 5015 -----VGEPLETDDAIVAKNFPDTPGAPGQPEAVETSEAITLQWTRTSDGAPIQ 5066
QY 4936 ---VE-----LTDGRHHIDQLDGTCSLLIAGLDRADAGCYTCQVSNK 4976
Db 5067 GYVIEKREVGSTWTKAAGNILDTKHRVTGLTKPKTYEFRVAVYNAAGQGEYS----- 5120
QY 4977 FGQVTHSNVVSSESEAESSSGGELDDAFRRARLRLFR-TKSPAESVDEELFLSA 5035
Db 5121 ---VNSVPIADNAPTRPKINMGMLTRDILAYAGERAKILVPPAASPA-----PKVTFSK 5172
QY 5036 DEGPAAPEEPADWOTYREDEHFI CIRFEALTQEARQVTRFOEMFATLIGVEIKLVEQ-- 5093
Db 5173 GENKISPTDPRVKEYSDPLATLTIKSELTDGGLYFVELENSQGSDDSIRLKVVDKPA 5232

Db 786 ---GGEKPM-TEAERQSLFPFGKVEKNDIPLPEKTVOQVDKICEWKCTYSRPAKI 840
Qy 1382 TWYKDGKLSSS---KYRMEAVGCTRRLLVVOQAQADTGEYSCEBAGGOLRSLFSLDVAEP 1438
Db 841 RMYKDKKIFSGGLKYKIVIEKNVCT--LIINNVEDDTGKVTCEANGVPTHQAULTVLEP 898
Qy 1439 KVVFAKEOPVHREVQAQAGASTTILSCVAQAQTEVWYKDKKLSFSSK---VRMEAVG- 1494
Db 899 PMKYSFLNPLPNTQBIYTKQAVLTCKVNTTPRALPVWYRGSAIQEGDPRFTIEKDAVGR 958
Qy 1495 CTRRLVVOQAQADAGEYSCEAGSQRLS-FHLHVAEPKAVFAKEQPASREVOAEAGTSAT 1553
Db 959 CT--LTIKEVEDDQAEWARTITQDVFSKVQVYVEEPHTFVPMKSQKVNESDL---AT 1013
Qy 1554 LSCEVAQAQTEVWYKDKKLS-SSSKVRMEAVGCTRRLLVVOQAQADAGEYSCKAGDOR 1612
Db 1014 LETDVNDKDAEVVWHDGKRIDIDGVKFKVSSNRKRLLIINGARIEDHGEYKCTTKDDR 1073
Qy 1613 LSFHLHVAEPKVVFAKEQ--PAHREVQAQAGASTILSCVAQAQTEVWYKDKKLS-- 1668
Db 1074 TMAQL-----IVDAKNKPIVALKDTEIEKDDVTLMQTKDTKTFGIFRNGKQISSMP 1127
Qy 1669 SSKVRMEAVGCTRRLLVVOQAQADAGEYSCEAGGOLRSLFHLHVALE--POISERPCR-- 1724
Db 1128 CGKFTQSRNGHTHLIKIKIEMNEADVTIEDQAGLRCNCNVTLBAEKRPILNWKPKKIE 1187
Qy 1725 ---REPLVVKEHEDIILTATLTPSAATVWLKOGVEIRRSRKH--ETASQGDTHLTAVH 1779
Db 1188 AKAGEPCVVVPFQIKGTRR-GDPKAI--LKNKPIDEMERKLVEIIVKDDVAEIVFK 1243
Qy 1780 GAQVLDSAIYSCRVG-----AEGQDF-----PV 1802
Db 1244 NPQLADTKWALEAGNSAGTALAPPELVFKDKPKPGKPLETKYNTAEGLDLVWGTPDPD 1303
Qy 1803 QVEEVAKFCRLLE-----PVCGLGGTVTLACELSPACAEVWVR-----CGNTQPRV 1850
Db 1304 EGAPKAYIIEMQERGSNWKAVGTGTDFVKOLKEH-GEYKFRVKNALNECGLSDPUT 1362
Qy 1851 GKRFQWVAEGPVRSITVLGLRAEDAGEYVCESRDDHTSAQLTVSVPRVVKFMSGLSTVVA 1910
Db 1363 GE--SVLAKNP-----YGVPGKPKNMDAIDVDKDHCT-----LAWEP 1398
Qy 1911 BEGGEATQCVVSPSDVAVVFRDCAALQPSK-----PAISQSGASHLSTISDLVLEDA 1965
Db 1399 BEDGAPTI-----TGYYIERREKSEKWHQVGTQPCDCELTKKKVEDK 1443
Qy 1966 GQI-TVEAEGASSAALRVREAPVLFK-KLEPQ-----TVEERSSVTLEVELT-R 2013
Db 1444 EYLVRKAVNKAAGPCDCHGKPIKNKAKKASPEFTGGIKDLRLKVGETIKYDVPISGE 1503
Qy 2014 PMPELRWTRNATALAPGRKNVEIHAEGARHRLVHLNVGFADRFF-----GCETPDD 2064
Db 1504 PLPECLVWVGKPLKAVGRVMSRSGKHIMKIENAVRADSGKFTILKNSSGSCD---- 1559
Qy 2065 KTOAKLTVMQVRLVRGLQVAEAREQGTATWE-----VOLSHADVDGS 2108
Db 1560 -STATVTVGRPTPPKGLDIAVCAQD-ATLSMNPDPDDGDPDLTGTYIVEAQDMDNK 1617
Qy 2109 WTRDGLRFQOGTCHLAVRGPMTLTLGLRPEDSGLMVFAEGVHTSARLVVTELPSVF 2168
Db 1618 YIEVG-----KVDNPTTTLUKVNGLR--NKGNYKFRKAVNNEGE----- 1654
Qy 2169 SRPLQ-DVVVTIEKEKVTLECELSRP-----NVDVRW---LKDG-----VELRAG 2208
Db 1655 SEPLSADQYTOIKDPWFDPGKREPTIDFADRIDIAWEPHKKOGGAPIEYIVVDRP 1714
Qy 2209 KTMATAOAGACRSLTITYCEPADQGVYVCDADHAQSSAVKVGQRTYTLIVRVLAEADG 2268
Db 1715 DTK-----EWKE-----VKRVPDTNASISGLKEGKEYQFRVAVNKAAGP 1754
Qy 2269 EIOFVAENAESPAQL-RVKELPVTILVRPLRDKIAEKGHGVLECOVSRASA----- 2318
Db 1755 Q-----PSESEKQLAKPKFIPAWL-----KHDNLKSIITVKAGATVRWEVKIGG 1798
Qy 2319 ---QVRWFKSQSELQPGKYELVSD-----GLYRKLIISDVHABD 2355
Db 1799 EPIPEVKFKGNQQLENG--IQLTIDTRKNEHTILCIPSAMRSDVGEYR-LTVKNSHGAD 1855
Qy 2356 E-----DIYTCADAGDKTSAQFFVEQSITIVRGLQDVTVMSP-----AP 2395
Db 1856 BEKANLTVLDPRSPKNGPLEVSDVF---EDNLNLSWKPPDDDDGGGPIEYVEYEKLDATG 1912
Qy 2396 AWFEC-----RPPKWLKGTVLQ-----AGNVGLE-----Q 2429
Db 1913 RWVPKAVKDKAHLIDGLKKGQTYQFRKAVNKEGASDALSTDKDTKAKNPYDEPKTGT 1972
Qy 2405 PSV-----RPPKWLKGTVLQ-----AGNVGLE-----Q 2429
Db 1973 PDVVWDADRVSLWEWEPKSGGAPITQYVIEKKGKGRDWOECVKSGSDOTNABILGLK 2032
Qy 2430 EGTVHRLMLRRTCSMTGTGPHFTVCKSRSSARLVVSDIPVLTTRPLEPKTGRELOSVL- 2488
Db 2033 EGEEYQFRKAV--NKAGP-----GEASDPKRVKAK-----PRNLKPIDREAMKTTITI 2080
Qy 2489 ---SCDF-----RPAPKAVQWYKDDTPLSPSEKFKMSLEGOMAEILRLIMPADAGVY 2538
Db 2081 KVGNDVEFDVPVRGEPPEPK-EMFNEKPVD-DQKIRIESEDYKTRFVLRGATRKHAGLY 2138
Qy 2539 RCOA---GSAHSTEVTVAREVTVTQPLQDAEATEBGWASFCELSHEDBEVMSL-- 2592
Db 2139 TLTATNAGSGDKHSVEVIVLGPSPLEVSNYE-----DRADLEWKVPE 2186
Qy 2593 --NGMPLYDSFHEISHK-----GRHRTLVLKSIQRADAGIVRASSILKVSTARS 2640
Db 2187 DDGGAPI---DHYEIEKMDLATGRWVPCGRSET-----TKTTPVNPQGHY 2230
Qy 2641 EVRVKPVFLKALDLSAE---ERGTALQCEVSDPEAHVVRKDGVOGL- 2687
Db 2231 KFRVAVNKEGSDPLTNTAILAKNPYEVFKVDKPEL-VDMDKDHDVLAWNAPDDGA 2289
Qy 2688 -----PSD-----KYDFLHTAGTRGLVHVDVSPED 2712
Db 2290 PIEAFVIEKKDKNGRWEALVVPDQKTATVPLNKEGEYQFRISARNKAGTGDPSPD 2349
Qy 2713 AGLTYCHVSGSETRARVPHDLHVGITKRLKTMVLEGESCFECVLSHESADAMTV 2772
Db 2350 ---RVAKPRNLAPRIH-----RELSDDTTVKVGTATLKFVHIDGEPAPD-VTWSF 2396
Qy 2773 GGTGVGSS-----SRF-----QATROGRKYILVVREAAPSDAGEVWFSVGRGLTSK- 2817
Db 2397 NGKGIGESKAOIENEPIYSRFPALPKALRKQSKYITATNINGTDSVTINIKVSKPTKP 2456
Qy 2818 ---ASLIVRERPAIILPLEDO-----WV----- 2838
Db 2457 KGPIEVTDFEDRATLDWKPPEDDGEPIEFVEIEKMTKDGIMVPCGRSGDTHFTVDSL 2516
Qy 2839 ---APGEDVELCEL-----SRAGTP-----VHM--- 2859
Db 2517 NKGDIKFRKAVNKEGSDPLETETDILAKNPFDRPGRPEPTDMSDHLKWDPP 2576
Qy 2860 LKDRKA-----IRKSQYD-----VVCETGMAMLVIRGASLKDAGEVTCV- 2900
Db 2577 LSDGAPITEEQIEKRTKYGWEPAITVPGGQTATV--PDLTPNEEVEFRVAVNKGCP 2634
Qy 2901 ---EASKSTASLHVEKANCFTBELTNLQVEEKGTAVFTCKTE-HPAATVTWR-KGLLE 2954
Db 2635 SDPSASKAVIAKPNLPHIDRDALKNLTIKAGQISFDVPVSGEPAPTVMHPDNR 2694
Qy 2955 LRASGKHOPSEGLTRLTISALEKADSDTTCIDGQASRAQLLVQGRVHIE----- 3009
Db 2695 IRNGRVLKNDPEYQSKLVVQMERGDSGTFTIKAVNANGEDATV---KINVIDKPTSP 2751
Qy 3010 ---DLEVDV-----VOEGSSAT----- 3023
Db 2752 NGPLDVSVDHGDVHTLWVRAPDDGGIPIENYVIEKYDTASGRWVPAKAVAGDKTTAVVD 2811

Db 4636 RPPKNGGDAITNVVVEKRTPGDWT--VGHVPGVTLRNLDRNDANTPYEVRRAENQVG 4693
Qy 4898 PMFSHTFGDTEAQVGDALRL-----ECVVASKADVRARWLK---DG 4935
Db 4694 -----VGEPLTDDAIVAKNPPDTPGAPQPEAVETSEEAITLQWTRPTSDG 4740
Qy 4936 -----VE-----LTDGRHHHIDQLGDGTCSLLIAGLDRADAGCYTC 4971
Db 4741 GAPIQGYVIERKEVSGSTWTKAAGNILDITKRVGTGLTPKTYEYFRVAAYNAAGOGES- 4799
Qy 4972 QVSNKFGQVTHSACVVSSESESSGELDDAFRAARELHRLFR--TKSPAESVDEE 5030
Db 4800 -----VNSVPTADNAPTRKINMGMLTRDILAYAGERAKILVPPFAASPA---PK 4846
Qy 5031 LPLSADEGPAEPEPADMQTYREDEHFICIRFEALTEARQAQVTRFQEMFATLIGIVEIKL 5090
Db 4847 VTFSGENKISPTDRVRKVEYSDFLATLTIEKSELTDGGLYFVELENSGSDSASIRLKV 4906
Qy 5091 VEQ--GPRRVMCISKETAP-----VVPPEPLSLLTSDAAPV--FLTELQNEQVQDG 5140
Db 4907 VDKPASQHIRV-----EDIAPDCCTLYWMPPS-----SDGSPITNVIVEXKLDLRHSDG 4956
Qy 5141 -----YVPSFDC-----VVTGQP- 5153
Db 4957 KWEKVSSFVRNLNYTVGGLIKDNRFRFRVRAETQYGVSEPCELADVVAKYQFVFNQPE 5016
Qy 5154 MPSV-----RWFKDGKLLJEEDDHYMINEDQ 5178
Db 5017 APTVRDKDSTWAELEWDPDROGSKIIGVOYQYRDTSSGRWINAKWDLSEQCHARVTGLR 5076
Qy 5179 QGQHOLIITAVVPADMGVIRCLAENSMGV-----STKAELR-----VDLTS-- 5220
Db 5077 QNGE-----FBRILIAKNAAGFKSPSPSERCOLKASRFGPGPIHVGAKSTG 5124
Qy 5221 -----TDVTDADATESSEYFSAQYLSRREGEQESTTDE----- 5256
Db 5125 RNHCITWMAPLEDGGSKITGVNIREYGSTLWTVASDY--NVREPEFTVDKLRFNDEYE 5183
Qy 5257 -----GOLPQVBEELRDLQVAPGTRLAKFOLKVKGYP 5289
Db 5184 FRVVAINAAGKIPSLPSGPIKIQESGSRPQIVVKPED-TAQPNRRRAVFTCEAVGRPE 5242
Qy 5290 PRLYFKDQOPITASAHIRMTGKILHTLEIISVTRDSGGVAAVYSNMGAAVSSARLL 5349
Db 5243 PTARLNRNGRELPESSRYRFEASDGVYKFTIKEVMDIDAGEYTVBSNPNYSGSDTATANLV 5302
Qy 5350 VRGDPDEEKPASDVHEQVLPRLMLERFTPKVKKGSSITFSPVKVEGRPVPTVHL---R 5406
Db 5303 VOAPPVIEK-----DVPNTILP-----SGDLVRLKIYPSG-TAPPRHSLVLNR 5344
Qy 5407 BEAERGVLWIGDTPGYTVASSAQHSLVLLDVGRHQHGYTTCIASNAAGQALCSALHV 5466
Db 5345 EE-----IDMDHPTIRIVEFDDHILITIPALSVREAGRYEYTVSNDSGEATTCFPLNV 5397
Qy 5467 SGLPKVBOEKVKEALISTFLOGTTOAISAOGLETASFADLGGQRKEEPLAAKEALHLS 5526
Db 5398 TGLPEAPO-----GPH 5409
Qy 5527 LAECTEEFLQKLTQITEMVSAKITQAKLVPGGDSDEDSKTPSPSRHGBSRSSSIQ 5586
Db 5410 ISNIGPSTATLSWRPPVTD-----GGSKITSYV 5438
Qy 5587 ESSSESDGDARGEIFDIYVVTADYLPGLAEQDAILREGQYVEVLDAHPLRMLVTRKP 5646
Db 5439 EKROLKSD-----EMVTVTSNVKONNYIVT--- 5463
Qy 5647 TKSSPSROGVSPAYLDRRLKLSPEWGAABAEFFPEGEAVSEDEYKARLSSVITQELLSSEQ 5706
Db 5464 -----GLFE-----NHEYEFVSA----- 5477
Qy 5707 AFVEELQFLSHLQHLERCHPVPIAVAGQKAVIFRNVDRDIGRPHSSFLQELQOQCDTDD 5766

Db 5478 -----QENENGIG----- 5484
Qy 5767 VAMCFIKNAFAEQYLEFLVGRVQAESVVVSTAIQEFYKKAABEALLAGDPSPQPPPIQ 5826
Db 5485 -----APLVS-----EHPHARLPDPPPTSPL- 5506
Qy 5827 HYLEQPERVQRYQALLKELIRNKARNQNCALLQAYAVVUSALPQRAENKLHVLSMENY 5886
Db 5507 -----NLBIVQVGGDYVTLWSQR----- 5524
Qy 5887 PCTLEALGEPTRQGHFIVWEGAPGARMKGNHRHVLFRNLHLVICKPRDSRTDTSVYV 5946
Db 5525 -----PLSDG-----GGRL-----RGYIV--EKQEEHEDEWFC 5551
Qy 5947 FRNMMLKSSIDINDQVEGDDRAFEVWQREDSVRKYLQARTAIKSSWVKEICGQQRL 6006
Db 5552 NQNPSPNNYVNLIDG-----RKY--RYRVFVANDAGLSDLAELDQTL 5594
Qy 6007 --ALPVWRPPPEEBELADCTAELGETVKLACRVCTGPKPVISWYKDGKAVQVDPHILLIE 6064
Db 5595 FQAGSGGPKIVSPSLDNEEVGRCVTFBCEISGSPREYRMPFKGCKELVDTSYKTLIN 5654
Qy 6065 DPDGSICALILDSLTGVDSGOYMCFAASAGNCSTLGLVQ-----VPPRPVNVKVRASP 6118
Db 5655 KGDQK-VLIINDLTSDDADEYTCRATNSGSTRSTRANIRIKTKPRVFPKTHGGYEAQ- 5712
Qy 6119 FVEGEDAQTCTIEGAPYQIRWYKDGALLTTGNKFTQLSEPRSGLLVLVIRAAKEDIG 6178
Db 5713 --KGETIELKIPYKAYPQGEARWTKDGEKIENNSKFSITDDK--PATLRISNASREDYG 5768
Qy 6179 LYECELNRLGASARASAEIRIQSPMLQAOEQCHREQLVAANVEDTTLERADQEVTSVLKRL 6238
Db 5769 EYRVVENSVDGSDGTNVTV----- 5789
Qy 6239 LGPRAPGSTGLTGTGPGPCRPAPALQETGSPPTGTSEAPAVPRVP--QPLHLEGF- 6295
Db 5790 -----ADVPE-PPRPPIIIEDEAVI 5810
Qy 6296 -EQPEAI---ARAQEWTVPIR-MEGAAMPGAGTGELLWDVHSHVVVRETTRTYTYQATD 6350
Db 5811 LSWKPPALDGGSLVNTYTIKREAGGSM-----SPCAKSRITYTTIE 5853
Qy 6351 -THTARPSMQVTTEDVQAQT---GGTAQFAIIEGDPQSPVTVTKDSVOLVDSTRLSQ 6406
Db 5854 GLRAGKQYEFRIIAENKHGQSKPCBPTA--PVLIPGDERKRRGYD-----VDEQCKIVR 5906
Qy 6407 QEGTYSVLVRHVASKAGVYTCLAQNTGGQVLCKAELLVLGGDNEPDSEKSHRRKLHS 6466
Db 5907 KGTVSS-----NYDNYVFDIWKQY-----POPVEIKHDH---VLD 5940
Qy 6467 FYEVEKEIGRGVFGVKKRVQHKGNKILCAAKFIPILRSRTRAQAVR-ERDILAAALSHPLVT 6525
Db 5941 HYDIEHELGTGAFGVVHRVTERATGNNFAAKFVMTPHESDKETVRKEIQTMSVLRHPTLV 6000
Qy 6526 GLLQOFETRKTLIILILECSSEELLDRLY-RKVVTVEAEVKYIQQVLVEGLHYLHSHGV 6584
Db 6001 NLHDAFEDDNEVMYIYEFMSGGELFEKVADEHNKMSDEAVEYMRQVCKGLCHMHENNVY 6060
Qy 6585 HLDIKPSNLMVHPAREDIKICDFGAQNTPAELQFSGYSPSPSPSPSPSPSPSPSPSPSP 6644
Db 6061 HLDLKPENIMTTTKRSNELKLIDFGLTAHLDPKQSVKVTGTTGAEPAAEVABKGPVGYT 6120
Qy 6645 DIWANGVISYLSLTSSPAGESRATLLNLEGRVSWSSPMAAH--LSEDAKOFIKATL 6702
Db 6121 DMWSGVGLSYLLSGLSPFGGENDDETLRNVKS--CDNMDDSAFSGISEDGDKDFIRKLL 6178
Qy 6703 QRAPOARSPAAQCLSHPWFLK-SMPABEAHINTKQLK 6739
Db 6179 LADPNTRMTIHQALEHPMLTPGNAPGRDQSPSSRYTK 6216

protein unc-22 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: A88852
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biological processes
A:Reference number: A75000; PMID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.elegans/
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and Science 283, 2103, 1999
A:Accession: A88852
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6831 <STO>
C:Cross-references: UNIPROT:Q23550; GB:chr_IV; PIDN:CAA98081.1; PID:g3881830; GSPDB:GN000000000
A:Gene: unc-22
A:Map position: 4
C:Superfamily: twitchin; fibronectin type III repeat homology; immunoglobulin homology;
Query Match 3.9%; Score 1619; DB 2; Length 6831;
Best Local Similarity 18.6%; Pred. No. 3.7e-39;
Matches 1442; Conservative 933; Mismatches 2680; Indels 2712; Gaps 325;
QY 120 RVREGSEATFCRVGSGPRPAYSWSKDGRRLLGEPDGPVRV-VEELGEA---SALRIRAR 175
DB 7 KCKQG-DLLMECHLEADPOPTIAQHSGNLL-EPGSRVVQTLPLGSLYKATLVKEPN 64
QY 176 PRDGGTYEVRANPIGASAAALVVDSDAATASRGTSTAAHLAHLORREARAGA 235
DB 65 AGDGAYKCTARNQIGESNANTLNLFAGAGGDEAKSRGPFV-----GK 108
QY 236 PASPSTGTTRTCTVTEGKHARLSVVTGPKETWKKDG-OLVTEGRRHVVED-AOEN 293
DB 109 PRILPFDG-----GALIVMECKVKSASTVAKMKDGVPLSMGLYHALFSLDGDQT 160
QY 294 FV-LKILFCKSDRLGYTCTASNLVGYTSVSVLVVREPAVPPFKRLQDLEVRKESAT 351
DB 161 YLCQLEIRGSSDAGQYRCNIRDQGETNANLNFEPP-----DPSEQRKRST 212
QY 352 FLCEVPOQSTEAWFKEETRLWASAKYIE-BEGTERLTVRNVSADDDAVICETPGS 410
DB 213 ---ASPRFSSRGPGSRP-----SSPKSMKSRGTPKR-----TLKPREGS 250
QY 411 RTVAELAVQGNLLRKLPRKTAVRVGDMTAMFCVELAVPVPVHMLRNOBEVWAGRVATSA 470
DB 251 PS-----KKLRSSTSPV-----NEEV-----SQS 270
QY 471 EGTRHTLTIQCCLEDVGQVAFMAGDCQSTSTRFCVSA PRKPPLOPPVDPVVKARMESVI 530
DB 271 ESRSSRT-----DKMEVDQVSG-----ASKRP----- 294
QY 531 LSWSPPHGERPVTIDGVLVEKKLGTYTWRICHEAENVAPPELTADVABEGNPQPRVS 590
DB 295 -DGLPPPG-----DEKKUR-----AGSP----- 312
QY 591 ALNSFGQSPYLEFPFGTVHAPKLAVRTPLKAVQAVEGGEVTFPSVLDLTVASAGWFLDQQA 650
DB 313 ---STRKSP-----SRKSASPTSRKSSAGG-----AASGTTGASASA 348
QY 651 LKASVYIEHCDTRHTTITREVPASLHGAQLKFVANGIESIRMEVRAAPGLTANKPPA 710
DB 349 TSATSGGSASSDASDKYT-----RDP----- 370
QY 711 AAAREVLARLHEEAQLAELSDQAAAVTWLKDGRITLSPGPKVEQASAGRRLVLRDVAR 770
DB 371 -----IVLEASR 377
QY 771 DDAGLYECVSRGGRIAYQLSVQGLARFLHKDMAGSCVDVAGGPAQFECETSEAHVHVHW 830
DB 378 -----SQTGRI-----GGSVVLEV-----QWCHSSTI-----IEW 403
QY 831 YKDGMEHGHSGERFLOEDVGTGTHRLVAATVTRQDEGTYSRV-----GEDSVDFRLRVS 884

DB 404 YRDGTLVRNSSE-YSQSPNGSIAKLQVNKLTEKSGLYKCHAKCDYGGQSSAMVYKIQS 462
QY 885 EPKVVFQAEQI-ARRKLOAEAGASATLSCEVAQAQTEVTWYKDGKLLSSSKVCMVEATGC 943
DB 463 DVEEELMKHRKDAEDEYQKEQKQTL-----QAEATKRVARRSKSKSKS----- 507
QY 944 TRRLVVOQAGQADAGEYSCEAGGQRLSPHLVDKPEKPVVFAKDQVAH---SEVOABAGANA 1000
DB 508 -----PAPQAKKSTTS-ESGRQBAS-----EVEHKSSSVRPDPDEES 544
QY 1001 TISCEVAQAQAEVVMYKDGKLLSSSLKVVHVEAKGCRRLVVQAGKCTDAGDYSCEARGQR 1060
DB 545 QLD-EIP-----SSGLTIPEE---RRRELLGQVGSDD----- 573
QY 1061 VSFRLHITPEKPMFAKQSVNEVQAEAGASAMLSCEVAQAQTEVTWYKDG----- 1111
DB 574 -----EVSEISBELPSFAGGPKRRKTDDK 597
QY 1112 -KLLS-----SSSKVGMVKGCTRLRLVLPQACKADAGEYSCEAGGQSVFHLHITPEKGVF 1166
DB 598 PKKUSIAPVSTNKSSDDEPSTPRR-----RSSIDMRRESVOEILEKTSPLVPSGASGA 652
QY 1167 AKQSVHNEVQAEAGTTAMLSCVA-QPQTEVTWYKDGKLLSSSKSVR---MEVRCGCTRL 1223
DB 653 PKIVEPENVTWENETAILTKVSGSPAPTRFRWFKGSRREVISGGRFKHITDGKEHTVAL 712
QY 1224 VVOQVGKADAGEYSCA-----GGQVVSFOLHITPEKAVFAKEQLVHNEVTRTEAGSATLS 1279
DB 713 ALLKCRSQDEGPTLTITENVHGTDSDVKLLVTSNGLDFRAMLKHRH---SQAG----- 764
QY 1280 CEVAQAQTEVTWYKDGKLLSSSKSVRIEAGCMRQLVVQAGQADAGEYTCAGGQRLSF 1339
DB 765 -----PQKDG-----EGGAGGG-----GGEKKPM 784
QY 1340 HLDVSEPKAVPAKEQ-----LAHRKVQABAGAIATLSCEVAQAQTEVTWYKDGKLLSS 1392
DB 785 -TEAERQSLPPGKKVEKWDIPLPEKTVQQVDKICEWKCTYSRPNKIRWYKDKEIFS 843
QY 1393 SS---KYRMEAVGCTRLRLVVOQAGQADAGEYSCEAGGQRLSFSLDVAEPKVVFAKEOPVH 1449
DB 844 GGLKIKVIEKNVCT--LIINNPEVDDTKYTCANGVPTHAQTLVLEPPMKYSFLNPLP 901
QY 1450 REVQAGASATLSCEVAQAQTEVMYKDGKLLSSSK---VRMEAVG-CTRLRLVVOQAG 1505
DB 902 NTQEIYRTKQAVLTCKVNTPRAPLVWYRGSKAIQEGDPRFIEKDAVGRCT--LTIRKEVE 959
QY 1506 QADAGEYSCAGSQRSL- FHLHVAEPKAVFAKEQAPASREVQAEAGTATLSCEVAQAQTE 1564
DB 960 EDDQAEWTARITQDVFSKVQVYVEEPRHTFVVPKMSQKVNESDL---ATLETVDNDKDAE 1016
QY 1565 VTWYKDGKLLS-SSSKVMEAVGCTRLRLVVOEAGQADAGEYSCKAGQORLSFHLHVAEPK 1623
DB 1017 VVMHDKGRIDIDGKPKVSSNNRKLIIINGARIEDHGEYKCTTKDRTWAQI----- 1070
QY 1624 VVFAKEQ--PAHREVQAEAGASATLSCEVAQAQTEVTWYKDGKLLSS---SSKVRVEAVGC 1679
DB 1071 IVDAKNKFIWALKDTEVIEKDDVTLMQTKDTKTPGIWFRNGKQISSMPGCKFETQSRNG 1130
QY 1680 TRRLVVOQAGQADAGEYSCAGGQRLSFRHLHVAELE--POISERPCR-----REPLVYKE 1732
DB 1131 THTLKIGKIEMNEADVVEIDQAGLRGSCNVTVLEAEKRPILNWKPKKIEAKAGEPCVVVKV 1190
QY 1733 HEDTILTLATLATSATVTWLDQGVETRRSRKH--ETASQDTHLTTLVHGAQVLDLSAIVS 1790
DB 1191 PFQIKGTRR-GDPKQAI---LKNKGPIDEMRKLVEVIIKDDVAEIVFKNPQLADTKNA 1246
QY 1791 CRVG-----ARGQDF-----PVQVEEVAAKFCR 1813
DB 1247 LELGNSAGTALAPFELPVKDKPKPKPLETKNVTASGLDLVWGTPDPDEGAPKAVYIIE 1306
QY 1814 LLE-----PVCBELGGTTLACELSPACAEVVR-----CGNTQPRVGRKRFQWVAEGP 1861

Db 1307 MOEGRSGNWKVGETGTGDFKVKDLKEH-GEYKFRVKALNECGLSDPLTGE---SVLAKNP 1363
Qy 1862 VRSLVTLGLRABADAGEYVCESRDDHTSAQLTVSVPRVVKFMSGLSTVVAEEGEGATFOCV 1921
Db 1364 -----YGVPGKPKNDKIDVDKDHCT-----LAWPEEDGGAPI--- 1398
Qy 1922 VSPSDVAVVFRDGLLQPSK-----FAISOGASHSLTISDLVLEDAGQI-TVEABGA 1975
Db 1399 -----TGIIERREKSEKMDHQVGTQPCCELTDKKVVEDKEYLYRVKAVNK 1446
Qy 1976 SSSAALRVREAPVLFK-KKLEPQ-----TVBERSVTLVELT-RPWPELRWTRNA 2024
Db 1447 AGPGCDHGKPIKMKAKKASPEFTGGIKDLRLKVGETIKYDVDPISGEPECECLVWVG 1506
Qy 2025 TALAPGKNVHHAEGARHLVLHNVGFADRGFF-----GCETPDDKTOAKLTVMR 2075
Db 1507 KPLKAVGRVMSSEKHKIMKIENAVRADSGKFTITLKNSSGSCD-----STATVTVGR 1561
Qy 2076 QVRLVRGLQAVAREGGTATME-----VOLSHADVDSWTRDGLRFQOG 2119
Db 1562 PTPPKGPLDIAOVCADG-ATLSWNPDDDDGGDPLTGYIVEAQDMONKGYIEVG----- 1614
Qy 2120 PTHCLAVRGMHTLTLGLRPEDSDGLMVPKABGVHTSARLVVTELPVSPSRPQ-DVWTT 2178
Db 1615 -----KVDPNTTLKVNGLR--NKGNYKFRVKAANNEGE-----SEPLSADQYTO 1657
Qy 2179 EKEKVTLECELSRP-----NVDVRW--LKDG-----VELRACKTMAIAAOGAC 2219
Db 1658 IKDPWDEPKGCRPEITDPADRIDIAWEPHPKDGAPIEEYIVVRDPDTK----- 1709
Qy 2220 RSLTYRCEFADQGVYVCDHAQAQSASVKVQGRYTYLLYRVLAEADAGEIOFVAENAES 2279
Db 1710 -----EWKE-----VKRVPDTNASISGLKEKEYQFRVRAVNKAGPGQ-----PSEPS 1752
Qy 2280 RAQL-RVKELPVTLRPLRDKIAMKHGVLBCQVSRASA-----QVRWPKG 2325
Db 1753 EKQLAKPFIAPWL-----XHDNLKSIITVKGATVWEVKIGGEPIPEVKNFKG 1801
Qy 2326 SQELOQPKYELVSD-----GLYRKLIISDVHADE-----DT 2358
Db 1802 NQOLENG--IQLTIDTRKNEHTILCPSAMRSDVGEYR-LTVKNGHGADEEKANLTVLDR 1858
Qy 2359 YTCADGVKTSQAFFVEQSITIVRGLQDVTWEP-----APAWFEC----- 2400
Db 1859 PSKNGPLEVSDVF---EDNLNLSWKPPDDDGGEPIEYVEVEKLDATGRWVPCAKVKOT 1915
Qy 2401 -----ETSIPIPV----- 2407
Db 1916 KAHIDGLKKGQTYQFRVKAVNKEGASDALSTDKDTKAKNPYDEPGKTGTPDVVDWADRV 1975
Qy 2408 -----RPPKWLKGTVLQ-----AGGNVGL-----QEGTVHRLMLRR 2440
Db 1976 SLEWEPKSDGGAPITQYVIEKKGHGRDQWCEGKVGSGDQTNAEILGLKEGEYQFRVKA 2035
Qy 2441 TCSTWGPVHFTVGSRSARLVVSDIPVVLTRLEPKTGRLOSVVL-----SCDF--- 2492
Db 2036 V--NKAGP-----GEASDPSRKRVAK-----PRNLKPMIDREAMKTIITIKVGNDFEFDVP 2083
Qy 2493 ---RPAPKAVOMKDDTPLSPSEKFKMSLEGQMAELRLRLMPADAGVYRCA-----GSA 2545
Db 2084 VRGEPPKK-EWIFNEKPDV-DQKRIESEDYKTRFVLRGATRKHAGLTYLTATNASGD 2141
Qy 2546 HSSTEVTVEARVTTGPIQDAEATEEGWASFSCELSHEDEBEVWSL-----NGMPLYNDS 2601
Db 2142 KHSVEVILGKSPSPGLPLEVSNVYE-----DRADLEWKVPEDDGAPI---D 2186
Qy 2602 FHEISHK-----GRHTLVILKSIORADAGIVRASSLKYSTSARLEVVKPVFLK 2651
Db 2187 HYEIEKMDIATGRWVPCGRSET-----TKTVPNLQPCHEYKFRVRAVNKEG 2233
Qy 2652 ALDDLISAE-----ERGTALQCEVSDPEAHVVRKDGVLG----- 2687
Db 2234 ESDPLTTNTAILAKNPYEVFGVKDPKPEL--VDMDKOHVDLAWNAPDDGGAPIEAFVIEKND 2292

Qy 2688 -----PSD-----KYDFLHTAGTRGLVVDVSPEDAGLYTCHVGSE 2723
Db 2293 KNGRWEELVVPDOKTATVFNLKEGEYQFISARNKAGTGDPSDPS-----RVVAK 2346
Qy 2724 ETRARVRVHDLHVGITKRLKTMVLEGESESCPECVLSHESADPAMWTVGKTKVSGS--- 2780
Db 2347 PRNLAPRIH-----REDLSOTTVKVGATLKFIVHIDGEPAD-VTWSFNGKIGESKQAQ 2399
Qy 2781 -----SRP-----QATRQGRKIYILVREAAPSDAGEVVFSVRGLTSK-----ASLIYR 2823
Db 2400 IENEPYISKRFALPKALRKQSGKYTITATNINGDSTVINIKVKSKPTPKGPIEVDVPE 2459
Qy 2824 ERPAALIKPLEDO-----WV----- 2838
Db 2460 DRATLDWKPEDDGGGEPIEFYIEKMNKTDGIVPCGRSGDHTFTVDSLNGKDHKFRVK 2519
Qy 2839 -----APGEDVELRCEL-----SRAGTP-----VHW-----LKDRKA----- 2865
Db 2520 AVNSEGSPDPLETETDILAKNPFDRPRGPEPTDWDSDHVDLKWDPPLSDGGAPIEY 2579
Qy 2866 -IRKSQYD-----VVCBGTWAMLVIRGASLKADAGEYTCV-----EASKSTA 2907
Db 2580 QIEKRTKYGRNEPAITVPGGOTTATV--PDLTPNEEYEFRRVAVNKGSPDPSDASKAVI 2637
Qy 2908 SLHVEEKANCFTTELTNLQVEEKGTAVFTCKTE-HPAATVTWTR-KGLLELBRASGHQPSQ 2965
Db 2638 AKPRNLKPHIDRDALKNLTIRAGQISFDVPVSGEPAPVTVTHWPNDRIRNGRVRKLDN 2697
Qy 2966 EGLTLRLITISALEKADSDTYTCDIQQAQSRQALLVQGRVHIE-----DLEDVD- 3015
Db 2698 PEYQSKLVVQMERGDSGTFTIKAVNANGEDEATV---KINVIDKPTSPNGPLDVSVDHG 2754
Qy 3016 -----VQEGSAT-----FRCR 3027
Db 2755 DHVTLNWRAPDDGGIPIENVVIEKYDTASGRWVPAKAVAGDKTTAVVDGLIPGHEYKFR 2814
Qy 3028 ISPANYE---PVHWF---LOKTPL-----HANELNEID-----AOPG--- 3059
Db 2815 VAAVNAEGESDPLEFTGTTLAKDPFKPKTNAPEITDMDKHVDLEWKPPANDGGAPTE 2874
Qy 3060 -----YHVLTLRLQA---LKDSGTIYF-----EAG---DQRASAA 3088
Db 2875 EYVEMKDEFFWMDVAHVPAQNTATVGNLKGSKYEFRIRAKNKGALGDPDSASAV 2934
Qy 3089 LRVTSEKPSVFSR-ELTDATITEGEDTLTVCTSTCDIP-MCWTQDKGTLRGSARCOLSHE 3146
Db 2935 AKARNVPPVIDRNSIQEIKVKAGQDFSLNIPVSGEPTPTITWTPESTPVESDDRMKLANE 2994
Qy 3147 GHRAQLLITGATLQDSGRY--KCEAGGACSSSIVRVHA-----RPVRFQBALKOLEV 3196
Db 2995 DGKTFHVKRALRSDTGTYYIIKAENENGTDTAEVKVTVLDHPSPRGPDLVDVNIKDGCD 3054
Qy 3197 L-----EGGAATLRCVLSV-AAPVKW-----CYGNN-----VLR 3225
Db 3055 LAWKEPEDDGGABISHYVIEKODAAATGRWTAGCESKOTNFHVDLDTQGHYEYKFRVAVNR 3114
Qy 3226 PGDKTSLRQEGAML-----ELV-----VRNLRPQDSGRYSCSFGDQTT 3263
Db 3115 HCDSDPLEAREAIIAKDPFDRADKPGTPEIVDMDKHADLKWTPPADDDGAPIE-GYLVB 3173
Qy 3264 SATLTVTALPAQFICKLRNKEKATEGATLRCLELSTKPAVEWRKGETLRDGRYCLROD 3323
Db 3174 MRTPSGDWVPAVTG-----AGE-LTATV-----DGLKPGQTY----- 3205
Qy 3324 GAMCELOIRGLAWDAABEYSCVCGEERTSASITIRPMPAHPIFGLRLH-----QESTEGAT 3378
Db 3206 -----QFRVKNALNAGE-----STPDDPSRTMWAKPRHLAPKIRNDRMFVQRVKAGOT 3253
Qy 3379 AT--LRCELSKAAPVW--RKGRESLRDGRHSLRQOGAVCELIQICGLAVADAGEYSCVC- 3434
Db 3254 LNFVNVGEPAKPIEWFLNGSPSSGNGNTHDNNNTKLTTKSTARADSGKYKIVAT 3313

Db 5128 LEDGSKITGYNVEIREYXSTLWTVASYD-NVREPEFTVDKLRBNDFNVEFRVVAINAAGK 5186
Qy 5257 -----GOLPQVVEBELROLVAPGTRLAKFQKVKGYGAPRLYWFQGP 5300
Db 5187 GIPSLSPGPIKIQESGGSRPQIVKPED-TAQPNRRAVFTCEAVGRPEPTARMLNRE 5245
Qy 5301 LTASAHIRMTGKILHTLEISVTRFDSGOVAAVYSNAMGAAYSSARLLVRGDPPEKP 5360
Db 5246 LPESRYRPEASDGVYKTKIKEVMDIDAGEYTVESNPYSGDTATANLVVQAPPVIEK-- 5303
Qy 5361 ASDVHEQLVPPRMRLERFTPKVKKGSSITFSVKVEGRPVPTVHML---REBAERGVLWIG 5417
Db 5304 --DVNTILP-----SGDLVRLKITYFSG-TAPFRHSLVLRNEE-----ID 5340
Qy 5418 PDTGTYVASSAQHSLVLLDVGROHQGTTCIASNAAGQALCSASLHVSLGPKVEQEK 5477
Db 5341 MDHPTIRIVFDHILITIPALSREAGRYEYTVSNDSEATTGFWLNVNLTGLPEAPQ--- 5397
Qy 5478 VKEALISTFLOQTQAIQAQGLTASPADLGGORKEEPLAAKEALGHLSLAEVGTETFLQ 5537
Db 5398 -----GPHIHSNIGSTATL 5412
Qy 5538 KLTQITEMVSAKITQAKLOVPGSDSDSKTSPASPRHGRSRPSSSIQESSSESDGDA 5597
Db 5413 SWRPVPTD-----GGSKITSYVVEKRLSKD--- 5438
Qy 5598 RGEFDIVVTVADYLPGLABQDAITLRGQVVEVLDAHPRLWLVRKPTKSSPSRQGWV 5657
Db 5439 -----EWTVTSNVKDMNYIIVT----- 5455
Qy 5658 SPAYLDRRLKSPENGAEAPEFFGAEVSEDEYKARLSVIOELLSSSEQAFVEELQFLQS 5717
Db 5456 -----GLPE-----NHEYEFVSA----- 5469
Qy 5718 HHLHLRCPHVPVIAVAGQAVIFRNRDIGNRHFSSFLQELQCCDCTDDVAMCFIKNOA 5777
Db 5470 -----QENGIG----- 5476
Qy 5778 FEQYLEFLVGRVQAEVSVVSTAIQBYKYKABEALLAGDPGPPPPHLYLEQPVVQ 5837
Db 5477 -----APLVS-----EHPHIALRPDPPTSP----- 5498
Qy 5838 RYQALLKELIRKARNRONCALQEAQYAVVSALPQAEKHLVSLMENYPGTLEALGPPI 5897
Db 5499 -----NLEIVQGVGYVTLWSQR-----PL 5518
Qy 5898 RQGHFIVWEGAPGAMPKGNHRHVPFLRNHLVICKRRDRSDTDTVSVYFRNMKLSID 5957
Db 5519 SDG-----GGRL-----RGYIV---EKQEBEHDFWFRCNQNPSPNNYN 5554
Qy 5958 LNDQVEGDDRAFEVWQEREDSVRKYLQARTAIKSSWVKEICGQQRL--ALPVWRPDP 6015
Db 5555 VPNLIDG-----RKY--RYRVFVAVNDAGLSDLAELDQTLFQAGSGGEGPK 5597
Qy 6016 FEELADCTAELGETVKLACRTGTPKPVISWYKDGKAVQVDPHILLIEDPGSCALLILD 6075
Db 5598 IVSPLSDLNBEVGRCVTFCEISGSPREYRWFPGCKELVDTSKYTLINKGDKQ-VLIIN 5656
Qy 6076 SLTGVDSGQYMCFAASAAGNSTLGLVQ-----VPRFVFNKVRSPFFVEGEDAQPTC 6129
Db 5657 DLTSDDADEYTCRATNSSGTSRANLRIKTKPRVPIPPKHGGYEAQ--KGETIELKI 5713
Qy 6130 TIEGAPYQIRWYKDGALLTTGNKFTLSEPRSGLLVLIVRAASKEDGLYECELAVNRLG 6189
Db 5714 PVKAYPQCEARWTKDGEKIENNSKFSITTDK--FATLRISNASREDYGEYRVVSVNG 5771
Qy 6190 SARASAEILRIOSPMLQAQEQCHREQLVAIVEDTLERADQEVTVLKLGLGKAPGPGTG 6249
Db 5772 SDSGVFNVTV----- 5781
Qy 6250 DLTRGCPGCPGAPALQETGSPPTVTGSEAPVPRVP--QPLLHEGP--EOEPEAI--- 6302
Db 5782 -----ADVPE-PPRFPIIENILDEAVILSWKPPALDGG 5813

Qy 6303 ARAQEWTVPIR-MEGAAMPGAGTCELLWDVHSHVVVRETTQRTTYTYQAID-THTARPPSMQ 6360
Db 5814 SLVNTYTIKREANGGSW-----SPCAKSRYYTYTTIEGLRAGKQYBFR 5856
Qy 6361 VTIEDVOAQT--GGAQFEAIIEGDPOPSVTVYKDSVOLVDSRSLSQOQEGITVSLVR 6417
Db 5857 IIAENKQSQKPCPEPTA--PVLIPGDERKRRRGYD-----VDEQKIVRGKGTVSS---- 5905
Qy 6418 HVASKDAGVYTCLAQNTGGQVLCKAEILLVGLGDNEPSEKSHRRKLSHPYEVKEEIGRG 6477
Db 5906 --NYDNVYFDMWQYV-----PQPVEIKHDH--VLDHYDIHEELGTG 5943
Qy 6478 VFGFVKRVQHGKNIKLAKEIPLRSRTRAQAYR-ERDILAAALSHPLVTGLLDQPETRKT 6536
Db 5944 AFGVYHVRTERATGNNFAAKFVMTPHESDKETVKEIQTMSVLRHPTLVNLDHAPEDDNE 6003
Qy 6537 LILILELCSSEILLDRLY-RKGVVTEAEVKYVIOQLVEGLHLSHSHVLDIKPSNITLM 6595
Db 6004 MWMIYEPMSGGELEKVADEHNKMSDEAVYMQVCKGLCHMHENNVYHLDLKPENIMP 6063
Qy 6596 VHPAREDIKICDFGAQNTPAELQFQYQSPFVSPFVPEIIOQNPVSEASDIWANGVYSL 6655
Db 6064 TTKRSNELKLDGFLTAHLDPKQSVKVTGTGAEPAAPEVAEGKPVGYTDMWSVGLSYI 6123
Qy 6656 SLTSSSPAGESDRATLLNVLEGRVSVSSPMAAH--LSEDAKDFIKATLQAPQAPPSAA 6713
Db 6124 LLSGLSPFGGENDDETNRNKS--CDNMDDSAFSGISEDGKDFIRKULLADPNTRMTIH 6181
Qy 6714 QCLSHHPFLK-SMPAEBAHFINTKOLK 6739
Db 6182 QALEHPMLTPGNAPGRDSQIPSSRYTK 6208

RESULT 10

Tl3931

projectin - fruit fly (Drosophila melanogaster) (fragment)

C:Species: Drosophila melanogaster

C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004

C:Accession: Tl3931

R:Daley, J.; Southgate, R.; Ayme-Southgate, A.

J. Mol. Biol. 279, 201-210, 1998

A>Title: Structure of the Drosophila projectin protein: isoforms and implication for pro

A:Reference number: Z17815; MUID:98300339; PMID:9636710

A:Accession: Tl3931

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-6658 <DAL>

A:Cross-references: UNIPROT:O76281; EMBL:AF047475; NID:G3337430; PID:G3337431; PIDN:AAC2

C:Genetics:

A:Gene: projectin

A:Cross-references: FlyBase:FBgn0005666

A:Map position: 4

A>Note: intron positions not resolved (incomplete sequence)

C:Keywords: muscle

Query Match 3.4%; Score 1392; DB 2; Length 6658;

Best Local Similarity 18.4%; Pred. No. 1.3e-32;

Matches 1567; Conservative 1067; Mismatches 2994; Indels 2902; Gaps 382;

Qy 182 YEVRAPENPLGNASAAALVVDSDAADTASRPGTSTAALLAHLQRREARREGAPASPPS 241

Db 2 FRVAVNRAG-----PGEPSDKTSIIAKRCFVKPFVIGE----- 36

Qy 242 TGTCTCTTEGCHARLSCVYTGEPKPEVTKKDGQVTEGRRHVYVEDAQNFVLKILFC 301

Db 37 -GLKNVTVKKQTIIRFDIKYDGEPEPAATWKGTDNLKFDNQRICLDQLERNSSITIKS 95

Qy 302 KQSDRLGYTCTASNLVGGQTYSSVLVYVRE-----PAVPFKKQLQDLEVRKESATFLCRVP 357

Db 96 VRKDTGTYKLVLSNSSGTIESEAQVWLDRLPLPGGPPEPE-----BIR 139

Qy 358 QPSTEAAWFKETELWASAKYIEEBGTERLTVRNVSAODDDAVYICETPEGSRTVAELA 417

Db 3574 --SEPSDMDTMRPRVLPKIIITPLNEVRILKGLIFHTDIHFGEPAEATWTLNSNP 3631
Qy 4290 LOANEMNDITVRQGLTHLLTLHKVLEADAGTVSFHVGTCSSEAQLKAVKNTVVRGLENV 4349
Db 3632 LLSNDRSIT--SIGHSHVHTVNCQSDSGIYHL-----LLNSSGI 3672
Qy 4350 EALEGGEALF-----ECOLSOPEVAHAHTWLLDDEPVRTSENAEVVFFENGRLHLL 4401
Db 3673 D--EGSFLVWLDROPPEGPMEYEITANSVTISWKP----- 3708
Qy 4402 KNLRPQDSCRVTFPLAGDMWTSFLTVR-----GWRLEILBPLKNAAVRAGAQAFTCTL 4455
Db 3709 ----PKDN-----GGSSEISSYVIEKRDILTHGGW----- 3733
Qy 4456 SEAVP-VGEASWYINGAAVQPDSDWTVTADGSHQALLRLSAQPHHAGEVTFACRDVA 4514
Db 3734 ---VPAVNVSAKYNHAVVR-----LLEGTMYELVRMAENLOGRSDPITSQDPVAK 3783
Qy 4515 ARLTVGLPDPPEDAEVVAHSHSTVTLTSLWAAPMSDGGGLCGYRVVEKGA TGOW-RLCH 4573
Db 3784 SOYTVPGAPGK-ELTSDSKNHTIKWKQPIISNGGSPFIIGYDIERRDRDVTGRWIKING 3840
Qy 4574 ELVPGPECVVDGLAGETVYRERVAAGVPGAGEPVHLPQTVRLAEPKVPVPPQSPAPESR 4633
Db 3841 QVPVTAEOYDDRVTSNHQYQVRISAVNAAGNG-----KTSEPSAIFNARPLREKPR 3891
Qy 4634 -----QVAAEGDVSLLELVV-ABAGEVIWHKGMERIQGGRPEVVVSQROQMLV 4681
Db 3892 FYFDGLGKRIKVRAGEPVNLNIPISGAPTTIEWKRGDLDLEGRKISYETNERTLFR 3951
Qy 4682 IKGFTAEDQGEYHCGLAQSGICPAAAATFOVALSPASVDEAPQPSLPPBAQEGDLHLWE 4741
Db 3952 IDDSNRNRSGKY-----TWTAANFEGKDTADIEVIVVDKPS-PP---EGPLSY--- 3995
Qy 4742 ALAKRRMSREPTLDSIS---ELPEDGRS-----QRLP---QEABEV 4778
Db 3996 ----TETAPDHISLHWSPKDDGSGDITGYIIEFTBFGVDDMKPVPFGTCPTNFT 4046
Qy 4779 APDLSEG-----YSTADBLA-----RTGDADLSH 4802
Db 4047 VKNLVEGKKYVRIIRAENIYGALEKGPVLVKSPPDPGAPSOPTISATPNSANLEW 4106
Qy 4803 TSSDD-----ESRAG-----TPS---LVTYLKAGRGPTSPSLASKVGA 4837
Db 4107 HPPDDCGKPIGTIVERRERGGEWIKCNNYPTNTSYTVSNLRDARGAYEPRVLAVNEAG 4166
Qy 4838 PAAPSVKPOQOQEPLAA-----VPPLGLDLS TKDLGDPMDKAAVKIQAAFKGYK 4888
Db 4167 PGHPS-KP---SDPMTAEHQRYRDPDPPEPKPDRI TRNGVTLSWRPPRTDGKRIKGYV 4222
Qy 4889 RKEMKQOQEG-----PMFSHTF----- 4904
Db 4223 --EMRPKNKGDKWTVNDIPINSTVTVPSLKEGEYSFRVVAENEVGRSDPSKPSQITI 4280
Qy 4905 -----GDTEAQVGDALRLBECVVASKADVARWRLKDPGVELTDGRHHHIDLQ 4950
Db 4281 BEQNKPKCMELGKVRDVCRAAGDDFSIHVPYLAPPKNAFWYSDNNLDDNNRVH-KHLT 4339
Qy 4951 DGTCSLLIAGLDRADAGCYTCOVSNKFGQVTHSACVVVSGSESAESSGSELDDAFRA 5010
Db 4340 DDAASVVVKNKRSDSGOYRLQFKDTS-----FDTA 4371
Qy 5011 ARRLHRLPRTKSPAEV-----SDELEFL-----SADEGPA-----BPEPADWQTYRED 5054
Db 4372 TINRVLDRLPSPPTRLRADVFGSDSLTYNPPNDGGSALQNYLIEKKE-ARSTWSKV 4430
Qy 5055 EHFICIRFEALTEARQAVTRQEMFATIGIGVEIKLVEQGRPRRVMCISKETP--APVP 5112
Db 4431 SSFCTVLP---VRI RNVLNKEYDFRVI---AENKYGQSDPANTSEPI LARHPDIPNTP 4484
Qy 5113 PEPLPSLLTSDAAVFLTLQNOEQVQGYPVSFVDCVVTGQPMPSVRWFQDKLLEDDHY 5172

Db 4485 GIPHGIDSTEDSITIAWTKPKH-----DG-----GSPi-----TGYYIEK---R 4520
Qy 5173 MINEDQGGHQLIITAVVPADMGVYRCLAENSMGVSTSKAELRVLDLTSTDYDTAADATES 5232
Db 4521 LLSDDKW-----TKAVHALCPDLSCKIPNL--IENAEYEFV-----AAVNAAGQ 4563
Qy 5233 SSYPSAQGYLSSRBEQGETSTTBEGQLPQVVEELRDLOVAPGTRLAKFQLKVKGY--PAP 5290
Db 4564 SAYSGSDLIFCRRPPHAPKITSDLSIP-----DMVTIAGD---EFRIPTVPVHASPRP 4613
Qy 5291 RLWYFKDGOPLTASAHIRMTGKILHLTLEIISVTREDSGOYAAVISNAMGAAYSSARLLV 5350
Db 4614 TASHSLNGLEVIPEERIKFDSNDYASYNKSAKRDETSYITILTNNKSGDTSASCHVT 4673
Qy 5351 RGPDEPEKEP--ASDVHEQLVPPRMLERFTPKVKKGSSITFSVKVGRGPVPTVHMLREE 5408
Db 4674 VDRPLPQGGFLNAYDI-----TPDT-----CTLAWKTPDLDGSGSPITNYVEK 4716
Qy 5409 AERGVLWIGPDTPGYTVASSAQHSLVLDVGRHQHTYTCIASNAAGQALCSASLHV-- 5466
Db 4717 LDNSGSWV-----KISSFVRNTHYDVMGLEPHYKYNFRVRAENQYG---LSDPLDIE 4766
Qy 5467 -----SGLPKVEEQEKVEALIST-----FLOG-----TT 5491
Db 4767 PMVAKHOFTVPDBFPQPKVIDWDSGNVTLIWTRPLSDGSGRIQGYQIYEYRDLNDSSWNA 4826
Qy 5492 QAISAQOLETASFADLGQRKEEPLAAKEALGHLSLAEVGTETBFLOKLTQITEMWSAKI 5551
Db 4827 YDIYIKTKYQLYNLINGSYEFRIKAKNAAG---LSKPSSPSLRFKL----- 4871
Qy 5552 TQAKLQVPGGSDSDSKTSPASPRHGRSRPSSSIQSSSESESDGAR--GEIFD----- 5603
Db 4872 -KGKFTVP-----SPPGAPQVTRVGVKNYVDLKWELRDLGGSRITGYIERRDIG 4921
Qy 5604 -IYVVTADYILPGAQDAITLRE-GQY-----VEVILDAHP 5637
Db 4922 AVWYKNDYVNLDTYETVYVNNLIEMGDYEFVFAVNSAGRSEPSLCTWPIKVEVLGKKKP 4981
Qy 5638 LRWLVRTKPTKSPRSQGVSPAYLDRRL-----KLSP--EW--GAAPPEPPEGEAVSE 5687
Db 4982 -DMITRL-----QDKVAPFGKYDTLQCAASGRKPSPTARWLNRNGKELOMNGRMTCD 5031
Qy 5688 DE---YKARLSSV-----IOELSSBOAFVEELQFLOSHLHLEPCP----- 5727
Db 5032 SKDGVFRLHSNVTQDGDYTCAMNS-LGFVNTSGYLKIGS-PINRCPSELKLPEDG 5089
Qy 5728 --HVPIAVAGOK--AVIFRNVRDIGRHSFLOEQCDDDDVAMCFIKNQAAFEQYLE 5783
Db 5090 NSKIKIPIYSGDQPLTVILKKNEV-----ISDNDTTHV---KVNIFDDYVA 5133
Qy 5784 FLV-----GRVQAESVVVS-TAIQEFYKKAEEBALLAGDPSQPPPP----- 5823
Db 5134 IYIANIVKSDGGYQIIEFTNESGSATGEFY-----VHITGMPSPATGPMGISYINKNSC 5187
Qy 5824 -----PLQHYLEQPVVERVQRYOALLKELIRNKARNRQNCALLLEQYAVVVSAL 5870
Db 5188 MLNWRPPYDGLKVSHYVTERKDVSSPHWITVTSSTCKDTAFNVQ--GLIENOEYIFRVM 5245
Qy 5871 PQRAENKHLVSLMENYPG-TLEALGHPIROGHFIVWEGAPGARMKPKHNRHFLPRNHL 5929
Db 5246 -----AVNENGWGPPLBGL-NPIRAKOPIDPPSPAPV-----LKSURSEE 5284
Qy 5930 VICKP-----RRDSRTDTSYVYFRNNMKLSSIDLNDQVGGDDRAFE 5970
Db 5285 TLCNPEWKPPESDGAHIQGYWIDKKEVSGKHMAGVNAITCAANQINCINLIEGRQYEP 5344
Qy 5971 VWQEREDSVKYLLOARTAIKSSWKEICGIIQORLALPWRPPDPFEEL-ADCTAELGE 6029
Db 5345 IFAQN-----VAGLSAIVR-----FTSSODNRSTGGFASIDCETATHANCIO--NH 5389
Qy 6030 TVKLACRVGTGPKPVLISWYKDGKAVQVDPHHILIEDPDG---SCALILDSLT-----GV 6080
Db 5390 NAQFTCTIPGVPKPTISWYKGAREIS-----NGAGITCTQKVSPLPKYDVFG 5438

QY 32 QIVGNPTQVSWEKDQQPVTAGARFLAQ-----DGDLYRLTILDALGDSGQY 80

Db 2241 PGPIPPVRIESSSTVAEGQTLDLSCVVGQAHAQVTVYKRG-----GSLPARHQVRG--S 2294
Qy 1038 RLVVQOAGKTADGYSCEA-RGQVRSFRLHIT-----EPKMPFAKE 1077
Db 2295 RLYIFQASPADAGQVVCRASNGMEASITVVTGTQGANLAYPAGSTQPIRIE----- 2347
Qy 1078 QSVNEVQAEAGASAMLSCEV-AQAQTEVTVYKDGKLLSSSKVGMVKGCTRRLLVLPQA 1136
Db 2348 -----SSSQVAGQTLDLNCVVPQGSHAQVTVHKRG-----GSLPVTHQTHGSLRLY--QA 2397
Qy 1137 GKADAGEYSCBAGGORVSFHLHI---TEPK-----GVFAKEQSVHNEVQAEAGTTAMLS 1187
Db 2398 SPADSGEYVCRVLGSSVPLEASVLVTIEPAGSVPALGVTPVTRIESSSSQVAGETQDLN 2457
Qy 1188 CEVA-QPQTEVTVYKDGKLLSSSKVMEVKGCTRRLLVVOQVGKADAGEYSC-----EAGG 1242
Db 2458 CLVAGQAHAQVTVHKRGSLPA-----RHQVHGSRLRL--QVTPADSGEYVCRVVGSSGT 2511
Qy 1243 QRVSFQLHITPKAVFAKEQLVHNEVTEA-----CASATLSCVA-QAOTEVTVYKD 1294
Db 2512 QEASVLVTI-QORLSGSHSQGVAPVRIESSASLANGHTLDLNCVASQAPHTITWYKR 2570
Qy 1295 GKLLSSSKVRIEAGCMRLVVQAGQADAGEYTC-----AGQRLSFHLDV-----SE 1345
Db 2571 GGSLLPS-----RHQIVG---SRLRIQVTPADSGEYVCHVSNAGSRETSLIIVIOGSGSH 2624
Qy 1346 PKAVFAKEQLAHRKVQABAGAIATLSCVA-QAOTEVTVYKDGKLLSSSKVMEVAGCT 1404
Db 2625 VPSVSPPIRIESSPTVVEGQTLNLCVVARQPAIITWYKRGGSLLPS-----RHQTHGSH 2680
Qy 1405 RRLVVQACQADTGEYSCAGQRLSFSLDVAEPKVFEA-----KEOPVHREV 1452
Db 2681 LRL--HOMSVADSGEYVCRAN-----NIDALEASIVISVPSGSPAPGSSPIRIES 2733
Qy 1453 QAQ---AGASATLSCV-AQAQTEVTVYKDGKLLSFSSKVRMEAVGCTRRLLVVQAGQAD 1508
Db 2734 SSSHVAEGETLDLNCVVPQHAQVTVHKRGSLPSHHQTR-----GSLRLHLHVSAD 2787
Qy 1509 AGEYSC-----EAGSORLSFHLHVAEP-KAVFAKEQAPASREVOAEAGT 1550
Db 2788 SGEYVCRVMSGSGPLEASVLVTIEASG---SSAVHVPAGGAPPIRIEPPSSRV-AE-GQ 2842
Qy 1551 SATLSCV-AQAQTEVTVYKDGKLLSSSKVMEVAGCTRRLLVQEAQADAGEYSCV-- 1607
Db 2843 TLDLKCVPQHAQVTVHKRGSLPARKQVH-----GPLLRL--NQVSPADSGEYSCQVT 2896
Qy 1608 --AGDQRLSFHLV--ABPKVVFAR--EQPAHREVQAE--AGASATLSCV-AQAQTEV 1657
Db 2897 GSSGTLEASVLVTIEPSSPGPIPAQLAQPITYEASSSHVTEGQTLNLCVVPQHAQV 2956
Qy 1658 TWYKDGKLLSSSKVMEVAGCTRRLLVVQAGQADAGEYSCAGG-----QRLSFRLLHA 1712
Db 2957 TWYKRGSLPA-----RHQTHGSQLRL--HLVSPADSGEYVCRASGPGPEQAEFTVTV 3010
Qy 1713 ELE-----POISERCREPLVVKHEHDIILTATLTPSA-ATVTV-----LKQGV 1757
Db 3011 PSEGSYRLRSPVISIDP---PSSTVQOQDASFCLIHDAAPISLEWKTNRNOLEDNV 3067
Qy 1758 EI-----RRSKKHET-----ASQG-----DTHTLTVHGAQLVDSALYSCRVAEG 1797
Db 3068 HISPNGSIITIVGRTPSNHGYRCVASNAYGVAQSVNLSVHGPTVS-----VLPEG 3120
Qy 1798 QDFPVQVEEVAAKFCRLLEPVCBGLGGTVTTLACELSPACAEVWRWCGNTPRVKGRFOMV 1857
Db 3121 -----PVMVYKGAVTLEC-----VSAGEPRSSARWTRI 3149
Qy 1858 AEGPVR-----SLTVLGLRAEDAGEYCESRDDHTSALQTVSV-----P 1896
Db 3150 SSTPAKLEQRTYGLMDSHAVLQISAKPSDAGTVVCLAQNALGTAAQKQVEVIVDTGANAP 3209
Qy 1897 RVVFMGSLSTVAABEGGEATFCQVSPSPDVAV--WFRDGLALLOPSEKFAISQSGASHSL 1955
Db 3210 GAPQVQAEAEELTVEAGHTATLRCSATCSGAPTTHWSKLSRSLPQWHLR-----EGDTL 3263

Qy 1956 TISDLVLEDAQOITVEAGASSAALRVREAPLVFKKLEP--QTVBERSV-----TLEV 2009
Db 3264 IIPVAAQDSQOYICNATSPAGHA-----EATIIHVESPPYATTVPHEVASQAGETVOL 3318
Qy 2010 E-LTRPPEL--RWTRNATAL---APGKNVIEHAEGARHLVHLNVGFPADGFFGCEPDP 2063
Db 3319 QCLAHGTPPLTFQWSRVGSSLPGRATARNELLHFERAAPE-----DSGRVRCVTN 3369
Qy 2064 DKQAKUVTENRQVRLVRG-----LOAVEAREQGTATMEV--QLSHADVDS----- 2108
Db 3370 KVGSAEAPALQ-----LVQGPGLSPATSIAGSTPTVQVTPQLETKSGASVEPHCAVPS 3425
Qy 2109 -----WTRDGLRFQOGPTCHLAVRGPMTLTLISGLRPEDSGLMVFKAEG-----VHTSA 2157
Db 3426 DQGTQLRWFKEG---GOLPPGHSVQDG---VLRIONLDQSCQGTIYCOAHGPMWKAQASA 3479
Qy 2158 RLVVTELP---VSFSRRLQDVVTEKEKVTLEC-ELSRPNVDVRLWKDGVLELRAKTMAI 2213
Db 3480 QLVQALPSVLINIRTSVQTVV--GHAVEFELALGDPRQVTVWSKVGHLRPG---IV 3534
Qy 2214 AAQACRSLLTIYRCFADQGVYVCDADHAQSSASVKVQGRTYTLIYRRLVAEDAGEIQFV 2273
Db 3535 QSGGVR---IAHVELADAGQYRCTATNA----- 3560
Qy 2274 AENAESAQLRVKELPVTILVRPLDKIAMEKHGRLGVCQVRSASA-----QV 2320
Db 3561 AGTTQSHVLLVLVQALP-----QISMPQ-----EVRVPAGSAAVFPCIASGYPTPDI 3606
Qy 2321 RWFKGSQELQPGKVELVSDGLYKRLIISDVHAEDEDITYCDA---GDVKTSAQFFVEE 2376
Db 3607 SWSKLDGSLPDSRLR-----NNWMLPSPVPQDAGTYVCTATNRQKVKAPFAHLQVPE 3660
Qy 2377 QSI 2379
Db 3661 RVV 3663

RESULT 12

titin - rabbit (fragment)
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 09-Jul-2004
C:Accession: S20901; I46520
R:Labait, S.; Gautel, M.; Lakey, A.; Trinick, J.
EMBO J. 11, 1711-1716, 1992
A:Title: Towards a molecular understanding of titin.
A:Reference number: S20897; MUID:92258380; PMID:1582406
A:Accession: S20901
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-6805 <LAB>
A:Cross-references: UNIPROT:Q28733; EMBL:X64696
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, February 1992
R:Labait, S.; Barlow, D.P.; Gautel, M.; Gibson, T.; Holt, J.; Hsieh, C.L.; Francke, U.;
Nature 345, 273-276, 1990
A:Title: A regular pattern of two types of 100-residue motif in the sequence of titin.
A:Reference number: I46520; MUID:90238553; PMID:2129545
A:Accession: I46520
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 4235-5250 <LA2>
A:Cross-references: EMBL:X17329; NID:gl756; PIDN:CAA35207.1; PID:g930251
C:Superfamily: titin; fibronectin type III repeat homology; immunoglobulin homology; pro
C:Keywords: muscle

Query Match 2.5%; Score 1046.5; DB 2; Length 6805;

Best Local Similarity 18.5%; Pred. No. 1.3e-22;

Matches 1394; Conservative 968; Mismatches 2748; Indels 2411; Gaps 346;

Qy 17 KAFVVSVKDATLSQCIQVGNPTPOVSW-----EKDQPVTAGARFRLAQDGLVRL 67

Db 476 KLVVVRAGCPRLPAIVRGPAPKVTWRKGVNVRKQVDLVDTMAF-----L 525


```
Db      5874  TVKEDKLEA-----PELDLDSE-----LAKGIVVRAGSGARHIFPKGRPTPDI 5918
Qy      5742  RNVRDIGHFHS-----SFLQ-ELQOCDTDDVAMCFIKNOAAFEQYLFVLGRV-QA 5791
Db      5919  TWSREEGFTDKVQVKGWNTQLSIDNCDR-NDAGKVIIVKLENSSGTKTAFVTVKVLDT 5977
Qy      5792  ESVVSTAIQFYKYAYEALLAGDPSPPPPLQHYLEQPV-----ERVQYQALLKELI 5847
Db      5978  PGPPONTAVKVKK---DSAVLVME-----PIIDGAKVRNYVDIKREST 6020
Qy      5848  RNKARNRQNCALLEQAVAVSALPOR-----AENKLHVSLMENYPG 5888
Db      6021  R-----KAYANVSKCNKTKFKVENLTGEGAIYFRVWAENEFVGVPVEVD 6067
Qy      5889  TLEALGEPIROGHFIVWE-GAPGARMPKWKGNHRHVLFRNLHLVICKPRDRSDRTDTSVYF 5947
Db      6068  AVKAAEPSPPGKVTLDVDSQTSASLME-----KPEHDDGGSRLVGYV- 6110
Qy      5948  RNMKLSIDLNDQVEGDRAFEVWQEREDSV-----RKYLQARTAILIKSW 5995
Db      6111  -----VEMQPKTEKWSVVAESKVCNAVVTGLSSGHEYQFRVKAYNEKGKS 6156
Qy      5996  VKEICGQORLALPWRPPDFEEELADCTAELGETVKLACRVGTGPKPVIISWYKDGKAVQ 6055
Db      6157  DPRVLGVPV-TAKDUTQPSKLPFKRYSVQAGEDLKEIPVIGRPRPEIFVWKGDEPLR 6215
Qy      6056  VDPHILIEDPDGSCALIDSLTGVDSQYQYMCFAASAGNCS-TLQKILVQVPRFVNKV 6114
Db      6216  -QTRVNVVEATATSTILHIKSSKDDFCYKITITATNSAGTATENLSVIVLEKPGPGVPV 6274
Qy      6115  R-----ASFPV--EGEDAQFT--CTI-----EGAPYQIRWYKDGAL-----LTTG 6151
Db      6275  RFDEISADFVLSMEPPAYTGGCQTSNVIIVEKRDTTTTTHIVSATVARTIKVTKLKTG 6334
Qy      6152  NKFO--TLSEPRSGLL-----VLV-----IRAAKEDLGLVECELVNRL 6188
Db      6335  SEYOFRIYAENRYKSTSLDSKPVIVQVPKEPGPGPTVTSVRDQMLVQWHEPVDNG 6394
Qy      6189  GSARASAEIRIQSPMLQAOEQCHRELQV-----AAVEDDTTL-----ERADQEVTSVLK 6236
Db      6395  GSKVLGYHL-----EQEKNSILWVKVKNKTLIDTKFKTKTGLDGLLEYEFKVSAAE 6444
Qy      6237  RLLGPKAPGSTDLTGCPGCPRGAPALQETGSGQPVTGTSEAPVAPRVPQPLLHEGPE 6296
Db      6445  NIVG-----IASLAKCFNAPVA-----RDPCDPP----- 6468
Qy      6297  QEPBAIRAQSWTVPIRMEGAWPGAG--TGELL-----WDVHSHV-VRET----- 6339
Db      6469  GRPEAIVITRN-NVTLKWKPAYDGGSKITGYIVEKDLDPGRMKASFTNVLETFEFTVS 6527
Qy      6340  ---TQRTYTYQAIDHTTA-----RPPSMQVTIED-----VQAQTGGTA 6374
Db      6528  GLVEDQRVEFRVIARNAAGNLSEPSSESGAITARDEIDAPNASLDPKYKDIVVHAGETFF 6587
Qy      6375  QFEALIEGDPOPSVTWYKDSVOLVDST-RLSQOQEGTYSYVLVRHVASKDAGVYTCIAQN 6433
Db      6588  VLEADIRKPIPDVWLVMDKGLELETTARMEIKSTIQKTLTILVKDCIRTDGGQYVVLKLSN 6647
Qy      6434  TGGQVLKCAELLVLGGDNEPD 6454
Db      6648  VGGTKSLDITVKVLDPRGPPE 6668

RESULT 13
568235
myosin-light-chain kinase (EC 2.7.1.117), 210K, nonmuscle - chicken
N:Contains: myosin-light-chain kinase, 108K, smooth muscle; telokin
C:Species: Gallus gallus (chicken)
C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 09-Jul-2004
C:Accession: S68235; A37099; B44389; A44389; S28227; S78216; A35093; A25810; S11652
R:Waterston, D.M.; Collinge, M.; Lukas, T.J.; Van Eldik, L.J.; Birukov, K.G.; Stepanova,
FEBS Lett. 373, 217-220, 1995
```

```
A:Title: Multiple gene products are produced from a novel protein kinase transcription r
A:Reference number: S68235; MUID:96033976; PMID:7589469
A:Accession: S68235
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1906 <WAT>
A:Cross-references: UNIPROT:P11799; EMBL:X52876; NID:G992992; PIDN:CAA37056.1; PID:G99292
R:Shoemaker, M.O.; Lau, W.; Shattuck, R.L.; Kwiatkowski, A.P.; Matrisian, P.E.; Guerra-S
J. Cell Biol. 111, 1107-1125, 1990
A:Title: Use of DNA sequence and mutant analyses and antisense oligodeoxynucleotides to
activity.
A:Reference number: A37099; MUID:90361738; PMID:2202734
A:Accession: A37099
A:Molecule type: mRNA
A:Residues: 649-1906 <SHO>
A:Cross-references: EMBL:X52876
R:Collinge, M.; Matrisian, P.E.; Zimmer, W.E.; Shattuck, R.L.; Lukas, T.J.; Van Eldik, L
Mol. Cell. Biol. 12, 2359-2371, 1992
A:Title: Structure and expression of a calcium-binding protein gene contained within a c
A:Reference number: A44389; MUID:92236611; PMID:1373815
A:Accession: B44389
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1695-1906 <COL>
A:Cross-references: GB:M88284; NID:G212237; PIDN:AAB53767.1; PID:G212238
A:Accession: A44389
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1750-1906 <CO2>
A:Cross-references: GB:M88283; NID:G211371; PIDN:AAA48647.1; PID:G211372
R:Yoshikai, S.I.; Ikebe, M.
Arch. Biochem. Biophys. 299, 242-247, 1992
A:Title: Molecular cloning of the chicken gizzard telokin gene and cDNA.
A:Reference number: S28227; MUID:93073972; PMID:1444462
A:Accession: S28227
A:Molecule type: mRNA
A:Residues: 1750-1906 <YOS>
A:Cross-references: EMBL:M96655; NID:G212744; PIDN:AAA49083.1; PID:G212745
A:Accession: S78216
A:Molecule type: DNA
A:Residues: 1750-1906 <YOW>
A:Cross-references: EMBL:M96987
R:Olson, N.J.; Pearson, R.B.; Needleman, D.S.; Hurwitz, M.Y.; Kemp, B.E.; Means, A.R.
Proc. Natl. Acad. Sci. U.S.A. 87, 2284-2288, 1990
A:Title: Regulatory and structural motifs of chicken gizzard myosin light chain kinase.
A:Reference number: A35093; MUID:90192792; PMID:2315320
A:Accession: A35093
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 935-1438, 'Q', 1440-1906 <OLS>
A:Cross-references: GB:M31048; NID:G212660; PIDN:AAA49069.1; PID:G212661
R:Guerrero Jr., V.; Russo, M.A.; Olson, N.J.; Putkey, J.A.; Means, A.R.
Biochemistry 25, 8372-8381, 1986
A:Title: Domain organization of chicken gizzard myosin light chain kinase deduced from a
A:Reference number: A25810; MUID:87157587; PMID:3030394
A:Accession: A25810
A:Molecule type: mRNA
A:Residues: 1258-1438, 'Q', 1440-1906 <GUE>
C:Genetics:
C:Introns: 1735/3; 1779/1; 1819/1
C:Superfamily: myosin-light-chain kinase, nonmuscle; fibronectin type III repeat homolog
C:Keywords: alternative initiators; ATP; calmodulin binding; phosphoprotein; phosphotran
F:542-599/Domain: immunoglobulin homology <IMM1>
F:935-1906/Product: myosin-light-chain kinase, 108K, smooth muscle (from 5.5kb transcrip
F:1098-1158/Domain: immunoglobulin homology <IMM2>
F:1451-1708/Domain: protein kinase homology <KIN>
F:1459-1467/Region: protein kinase ATP-binding motif
F:1750-1906/Product: telokin (kinase-related protein KRP) (from 2.7 kb transkript) #stat
F:1808-1869/Domain: immunoglobulin homology <IMM3>
```

```
Query Match          2.5%; Score 1012; DB 1; Length 1906;
Best Local Similarity 19.7%; Pred. No. 3.4e-22;
Matches 488; Conservative 318; Mismatches 860; Indels 816; Gaps 76;
```


THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2005, 17:29:13 ; Search time 693 Seconds
(without alignments)
5887.802 Million cell updates/sec

Title: US-10-077-130-5

Perfect score: 41273

Sequence: 1 MDQPGSGAPRFLTRPKAFV.....RNREKRALLYKRNHLAQVR 7968

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	32134	77.9	6620	2 Q96AA2	Q96aa2 homo sapien
2	8486.5	20.6	2212	2 Q8NH3	Q8nh3 homo sapien
3	6902	16.7	1319	2 Q9HCD3	Q9hcd3 homo sapien
4	6619.5	16.0	1596	2 Q9HCL6	Q9hcl6 homo sapien
5	6128.5	14.8	1579	2 Q69SL2	Q69sl2 mus musculus
6	5358.5	13.0	1040	2 Q8NH2	Q8nh2 homo sapien
7	3865	9.4	767	2 Q86YC7	Q86yc7 homo sapien
8	3580.5	8.7	26926	2 Q8WZB3	Q8wzb3 homo sapien
9	3577.5	8.3	26926	2 Q10466	Q10466 homo sapien
10	3427	8.3	34350	2 Q8WZ42	Q8wz42 homo sapien
11	3389	8.2	658	2 Q8NH4	Q8nh4 homo sapien
12	3313	8.0	7962	2 Q10465	Q10465 homo sapien
13	3270	7.9	1020	2 Q8NH0	Q8nh0 homo sapien
14	3182.5	7.7	8081	2 Q7Z120	Q7z120 caenorhabdi
15	3113	7.5	646	2 Q8NH8	Q8nh8 homo sapien
16	3110	7.5	1944	2 Q69SL3	Q69sl3 brachydanio
17	2793	6.8	551	2 Q8NH7	Q8nh7 homo sapien
18	2653	6.4	542	2 Q8NH5	Q8nh5 homo sapien
19	2629	6.4	4162	2 Q8918	Q8918 gallus gall
20	2486	6.0	6632	1 UN89_CAEEL	Q01761 caenorhabdi
21	2291	5.6	4039	2 Q7Z246	Q7z246 brachydanio
22	2273	5.5	5516	2 Q7Z248	Q7z248 brachydanio
23	2233.5	5.4	3094	2 Q69SL1	Q69sl1 homo sapien
24	2203	5.3	3262	2 Q9EQJ5	Q9eqj5 mus musculus
25	2189.5	5.3	5604	2 Q8WZ53	Q8wz53 homo sapien
26	2168	5.3	3950	2 Q7YRF5	Q7yrf5 canis famli
27	2144	5.2	17903	2 Q7RTL4	Q7rtl4 drosophila
28	2087	5.1	16215	2 Q9NFS3	Q9nfs3 drosophila
29	2087	5.1	18074	2 Q917U4	Q917u4 drosophila
30	2072.5	5.0	5636	2 Q96RW7	Q96rw7 homo sapien
31	2067	5.0	2242	2 Q9P2P9	Q9p2p9 homo sapien

ALIGNMENTS

RESULT 1

ID	Q96AA2	PRELIMINARY;	PRT;	6620 AA.
AC	Q96AA2;			
DT	01-DEC-2001	(Tremblrel. 19, Created)		
DT	01-DEC-2001	(Tremblrel. 19, Last sequence update)		
DT	01-MAR-2004	(Tremblrel. 26, Last annotation update)		
DE	Obscurin.			
GN	Name=OBSCN;			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Cardiac;			
RX	MEDLINE=21342081; PubMed=11448995; DOI=10.1083/jcb.200102110;			
RA	Young P.W., Ehler E., Gautel M.;			
RT	"Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor			
RT	protein involved in sarcomere assembly.";			
RL	J. Cell Biol. 154:123-136(2001).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Cardiac;			
RA	Gautel M.S.;			
RL	Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: Contains 1 PH domain.			
DR	EMBL; AJ002535; CAC44768.1; -.			
DR	HSSP; P56276; 1FHG.			
DR	GeneW; HGNC:15719; OBSCN.			
DR	InterPro; IPR003961; FN_III.			
DR	InterPro; IPR008979; Gal_bind_like.			
DR	InterPro; IPR007110; Ig-like.			
DR	InterPro; IPR003598; Ig_c2.			
DR	InterPro; IPR000048; IQ_region.			
DR	InterPro; IPR001849; PH_region.			
DR	InterPro; IPR000219; RhoGEF.			
DR	InterPro; IPR001452; SH3.			
DR	InterPro; IPR001412; tRNA-synt_1.			
DR	Pfam; PF00041; fn3; 2.			
DR	Pfam; PF00047; ig; 15.			
DR	Pfam; PF00612; IQ; 1.			
DR	Pfam; PF00169; FH; 1.			
DR	Pfam; PF00621; RhoGEF; 1.			
DR	SMART; SM00060; FN3; 2.			
DR	SMART; SM00408; IGC2; 18.			
DR	SMART; SM00015; IQ; 1.			
DR	SMART; SM00233; PH; 1.			
DR	SMART; SM00326; SH3; 1.			
DR	PROSITE; PS00178; AA TRNA_LIGASE_I; UNKNOWN_1.			
DR	PROSITE; PS50010; DH_2; 1.			
DR	PROSITE; PS50853; FN3; 2.			
DR	PROSITE; PS50835; IG_LIKE; 46.			
DR	PROSITE; PS50096; IQ; 1.			

32	2067	5.0	2995	2	O696W0	O696w0 brachydanio
33	2048	5.0	5198	2	O76518	O76518 caenorhabdi
34	2043.5	5.0	5175	2	Q810L3	Q810l3 caenorhabdi
35	2035.5	4.9	3410	2	Q7TN00	Q7tn00 rattus norv
36	1996.5	4.8	803	2	Q80ZFS	Q80zfs rattus norv
37	1956.5	4.7	17352	2	Q95YM2	Q95ym2 procambarus
38	1896	4.6	4816	2	O8T103	O8t103 bombyx mori
39	1891.5	4.6	8625	2	Q86GD6	Q86gd6 procambarus
40	1837	4.5	4824	2	Q95YM1	Q95ym1 procambarus
41	1763.5	4.3	4796	2	Q9W055	Q9w055 drosophila
42	1736.5	4.2	4796	2	Q9NL88	Q9nl88 drosophila
43	1686.5	4.1	8930	2	O7KQP7	O7kqp7 drosophila
44	1680.5	4.1	8648	2	O7KQP6	O7kqp6 drosophila
45	1677	4.1	8943	2	Q9V4F7	Q9v4f7 drosophila

QY 2101 SHADVGSWTRDGLRFOQGPTCHLAVRGPMTHTLTLGLRPEDSGLMVFKABGVHTSARLV 2160
DB 2101 SHADVGSWTRDGLRFOQGPTCHLAVRGPMTHTLTLGLRPEDSGLMVFKABGVHTSARLV 2160
QY 2161 VTELVPVFSRPLQDVVTTEKEKVTLCELSRPNDVRWLKDGVELRAGKTWAIAGQACR 2220
DB 2161 VTELVPVFSRPLQDVVTTEKEKVTLCELSRPNDVRWLKDGVELRAGKTWAIAGQACR 2220
QY 2221 SLTIYRCBFADQGVVYCDHAHQSSASVKVQGRTYTLIYRRYLABDAGEIQFVAENAESR 2280
DB 2221 SLTIYRCBFADQGVVYCDHAHQSSASVKVQGRTYTLIYRRYLABDAGEIQFVAENAESR 2280
QY 2281 AOLRVKELPVTILVRPLRDKIAMEKHURGVLECOVSRASQVRWFKGSQELQPGPKYELVSD 2340
DB 2281 AOLRVKELPVTILVRPLRDKIAMEKHURGVLECOVSRASQVRWFKGSQELQPGPKYELVSD 2340
QY 2341 GLYRKLIISDVHAEDEDTYTCDAGDVKTSAQFFVEEQSIITIVRGLODVTVMEPAPAWPEC 2400
DB 2341 GLYRKLIISDVHAEDEDTYTCDAGDVKTSAQFFVEEQSIITIVRGLODVTVMEPAPAWPEC 2400
QY 2401 ETSIPSVRPPKWLKGTVLQAGNVGLBQEGTVHRLMLRRTCSWTGVPVHFTVVGKSRSSA 2460
DB 2401 ETSIPSVRPPKWLKGTVLQAGNVGLBQEGTVHRLMLRRTCSWTGVPVHFTVVGKSRSSA 2460
QY 2461 RLWVSDIPVILTRPLEPKTGRELOQSVLSCDFRPAPKAVQVYKDDTPLSPSEKFKWMSLEG 2520
DB 2461 RLWVSDIPVILTRPLEPKTGRELOQSVLSCDFRPAPKAVQVYKDDTPLSPSEKFKWMSLEG 2520
QY 2521 QMAELRIURLMPADAGVYRCQAGSAHSTEVTVAREVTVTGLODAEATEEGWASFSCE 2580
DB 2521 QMAELRIURLMPADAGVYRCQAGSAHSTEVTVAREVTVTGLODAEATEEGWASFSCE 2580
QY 2581 LSHEDDEEVSUNGMPLYNDSPFHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTARS 2640
DB 2581 LSHEDDEEVSUNGMPLYNDSPFHEISHKGRHRTLVLKSIQRADAGIVRASSLKVSTARS 2640
QY 2641 EVRKVPVFLKALDLSAERGTALQCEVSDPEAHVVRKDGVOQLGPSDKYDFLHTAGT 2700
DB 2641 EVRKVPVFLKALDLSAERGTALQCEVSDPEAHVVRKDGVOQLGPSDKYDFLHTAGT 2700
QY 2701 RGLVVHDVS PEDAGLYTCHVGSEETRARVRVHDHLHVGITKRLKTMVEVLEGESCSPECVLS 2760
DB 2701 RGLVVHDVS PEDAGLYTCHVGSEETRARVRVHDHLHVGITKRLKTMVEVLEGESCSPECVLS 2760
QY 2761 HESASDPAMWTVGGKTVGSSSRFQATROGRKYILVVRREAAPSDAGEVVFVSRGLTSKASL 2820
DB 2761 HESASDPAMWTVGGKTVGSSSRFQATROGRKYILVVRREAAPSDAGEVVFVSRGLTSKASL 2820
QY 2821 IVRERPAAI IKPLEDQWAPGEDVELRCELGRAGTPVHMLKDKRAIRKSQKYDVVCEGTM 2880
DB 2821 IVRERPAAI IKPLEDQWAPGEDVELRCELGRAGTPVHMLKDKRAIRKSQKYDVVCEGTM 2880
QY 2881 AMLVIRGASLKDAGTYTCEVBASKSTASLHVEEKANCFTBELTNLQVEEKGTAVFTCKTE 2940
DB 2881 AMLVIRGASLKDAGTYTCEVBASKSTASLHVEEKANCFTBELTNLQVEEKGTAVFTCKTE 2940
QY 2941 HPAATVTVWRKGLLELRASGKHQPSQEGTLRLTISALEKADSDTYTCDIGQAQSRQALLV 3000
DB 2941 HPAATVTVWRKGLLELRASGKHQPSQEGTLRLTISALEKADSDTYTCDIGQAQSRQALLV 3000
QY 3001 QGRRVHIITEDLVVDVQSGSATFCRIS PANYEPVHWFLOKTPHANELNEIDAQPGGY 3060
DB 3001 QGRRVHIITEDLVVDVQSGSATFCRIS PANYEPVHWFLOKTPHANELNEIDAQPGGY 3060
QY 3061 HVLTLRQALKXDSGTIYFEAGDORASALRVTEKPSVFSRELTDAITTEGEDTLVLCETS 3120
DB 3061 HVLTLRQALKXDSGTIYFEAGDORASALRVTEKPSVFSRELTDAITTEGEDTLVLCETS 3120
QY 3121 TCDIPMCTKDGKTURGSARCOLSHEGHRAQLLITGATLQDSGRYKCEAGGACSSSIVRV 3180
DB 3121 TCDIPMCTKDGKTURGSARCOLSHEGHRAQLLITGATLQDSGRYKCEAGGACSSSIVRV 3180

QY 3181 HARPVRFQEALKDLEVLGGGAATLRCVLSSVAA PVKWCYGNVLRPGDKYSILROEGAMLE 3240
DB 3181 HARPVRFQEALKDLEVLGGGAATLRCVLSSVAA PVKWCYGNVLRPGDKYSILROEGAMLE 3240
QY 3241 LVVRNLRPQDSGRYSYCSFGDQTTSATLTVTALPAQFIGKLRNKEATEGATATLRCELSKT 3300
DB 3241 LVVRNLRPQDSGRYSYCSFGDQTTSATLTVTALPAQFIGKLRNKEATEGATATLRCELSKT 3300
QY 3301 APVEWRKSGSETLRDGDYRCLRDQGMCELOIRGLAMVDAABEYSCVCGBERTSASITIRPM 3360
DB 3301 APVEWRKSGSETLRDGDYRCLRDQGMCELOIRGLAMVDAABEYSCVCGBERTSASITIRPM 3360
QY 3361 PAHFIGRLRHQESIEGATATLRCELSKAAPVEMWRKGRESLRDGRHSLRDQGAVALCELOIC 3420
DB 3361 PAHFIGRLRHQESIEGATATLRCELSKAAPVEMWRKGRESLRDGRHSLRDQGAVALCELOIC 3420
QY 3421 GLAVADAGEYSCVCGBERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLWCELSKVAPVE 3480
DB 3421 GLAVADAGEYSCVCGBERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLWCELSKVAPVE 3480
QY 3481 WRKGPENLRDGDYRILROEGTRCELOICGLAMADAGEYLCVCGQERTSATLTIRALPARP 3540
DB 3481 WRKGPENLRDGDYRILROEGTRCELOICGLAMADAGEYLCVCGQERTSATLTIRALPARP 3540
QY 3541 IEDVKNQEARREGATAVLQCELSNAAAPVEMWRKSGSETLRDGDYRSLRDGTCKELOIRGLAM 3600
DB 3541 IEDVKNQEARREGATAVLQCELSNAAAPVEMWRKSGSETLRDGDYRSLRDGTCKELOIRGLAM 3600
QY 3601 ADTGEYSCVCGQERTSAMLTVRALPIKFTTEGLRNEEAETEGATAVLRCELS KMAPVEMWKG 3660
DB 3601 ADTGEYSCVCGQERTSAMLTVRALPIKFTTEGLRNEEAETEGATAVLRCELS KMAPVEMWKG 3660
QY 3661 HETLRDGRHSLRDQGMARCELOIRGLVAEDAGEYLCMCGKERTSAMLTVRAMPSPKFI EGL 3720
DB 3661 HETLRDGRHSLRDQGMARCELOIRGLVAEDAGEYLCMCGKERTSAMLTVRAMPSPKFI EGL 3720
QY 3721 RNEEAETEGDTATLNCELSKAAPVEMWRKGHETLRDGRHSLRDGSRCELOIRGLAVVDAG 3780
DB 3721 RNEEAETEGDTATLNCELSKAAPVEMWRKGHETLRDGRHSLRDGSRCELOIRGLAVVDAG 3780
QY 3781 EYSCVCGQERTSATLTVRALPARFIEDVKNQEARREGATAVLQCELSKAAPVEMWRKSGSETL 3840
DB 3781 EYSCVCGQERTSATLTVRALPARFIEDVKNQEARREGATAVLQCELSKAAPVEMWRKSGSETL 3840
QY 3841 RGDYRSLRDGTRCELOIHGLSVADTGEYSCVCGQERTSATLTVRAPQPVFPREPQLSLQ 3900
DB 3841 RGDYRSLRDGTRCELOIHGLSVADTGEYSCVCGQERTSATLTVRAPQPVFPREPQLSLQ 3900
QY 3901 AEEGSTATLQCELSEPTATVVMKGGLOLQANGREPRELQCTAELVLQDLQREDTGEYT 3960
DB 3901 AEEGSTATLQCELSEPTATVVMKGGLOLQANGREPRELQCTAELVLQDLQREDTGEYT 3960
QY 3961 CTCSQATSATLTVTAAAPVRFELRLOHQEVDEGGTAHLCCELSRAGASVEMWRKSLQLFP 4020
DB 3961 CTCSQATSATLTVTAAAPVRFELRLOHQEVDEGGTAHLCCELSRAGASVEMWRKSLQLFP 4020
QY 4021 CAKYQMVQDGAABELLVRGVEQEDAGDYTCDTGHTQSMASLSVRVPRPKFKTRLQSLQBE 4080
DB 4021 CAKYQMVQDGAABELLVRGVEQEDAGDYTCDTGHTQSMASLSVRVPRPKFKTRLQSLQBE 4080
QY 4081 TGDITARLCCQLSDAESGAVVQWMLKEGVELHAGPKYEMRSQCATRELLTHOLEAKDGTGEYA 4140
DB 4081 TGDITARLCCQLSDAESGAVVQWMLKEGVELHAGPKYEMRSQCATRELLTHOLEAKDGTGEYA 4140
QY 4141 CVTGQKTAASLRVTEPEVTIVRGLVDAEVTADEDFECSVSRAGATGVOWCLOGLPLQ 4200
DB 4141 CVTGQKTAASLRVTEPEVTIVRGLVDAEVTADEDFECSVSRAGATGVOWCLOGLPLQ 4200
QY 4201 SNEVTEVAVRDRGRIHTLRLKGVTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPLQDVQ 4260
DB 4201 SNEVTEVAVRDRGRIHTLRLKGVTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPLQDVQ 4260
QY 4261 LSEGDQASFCQLRSASQGEARWALGGVPLQANEMNDITVEQGTLLHLTLHKVTLEDAGT 4320

Db	4261	LSBQDASFOCKLRSASQEARWALGGVPLQANEMKIDITVEQGTJHLTLTKHVTLEDAGT	4320
Qy	4321	VSFHVGTCSSEAQLKVTAKNTVVRGLENVEALEGEALFECQLSQPEVAHAHTWLLDDEPV	4380
Db	4321	VSFHVGTCSSEAQLKVTAKNTVVRGLENVEALEGEALFECQLSQPEVAHAHTWLLDDEPV	4380
Qy	4381	RTSENAEVFFENGRLHLLLLKNLRPDQSCRVTFLAGDMVTSFALTVRGWRLEILEPLKN	4440
Db	4381	RTSENAEVFFENGRLHLLLLKNLRPDQSCRVTFLAGDMVTSFALTVRGWRLEILEPLKN	4440
Qy	4441	AAVRAGAQRCTCTLSEAVPVGEASWYINGAAVOPDDSDTWTADGSHQALLRSAPHH	4500
Db	4441	AAVRAGAQRCTCTLSEAVPVGEASWYINGAAVOPDDSDTWTADGSHQALLRSAPHH	4500
Qy	4501	AGEVTFACRDVASARLTVLGLPPDPEDAEEVAHSSHVTVLSAAPMSDGGGJCGVRVE	4560
Db	4501	AGEVTFACRDVASARLTVLGLPPDPEDAEEVAHSSHVTVLSAAPMSDGGGJCGVRVE	4560
Qy	4561	VKEGATGQWRILCHELVPGCEVVDGLAPGETYRFRVAAVPGVAGEPVHLPTQVRLAEP	4620
Db	4561	VKEGATGQWRILCHELVPGCEVVDGLAPGETYRFRVAAVPGVAGEPVHLPTQVRLAEP	4620
Qy	4621	KPVPPQPSAPRSQVAAGEDVSLLEVVAEAGEVIMHKGMERTQPGGRFEVVSQGRQML	4680
Db	4621	KPVPPQPSAPRSQVAAGEDVSLLEVVAEAGEVIMHKGMERTQPGGRFEVVSQGRQML	4680
Qy	4681	VIKGFTAEQCEYHCGLAQSGISCPAAATFOVALSPASVDEAPQPSLPPEAAQSGDLHLW	4740
Db	4681	VIKGFTAEQCEYHCGLAQSGISCPAAATFOVALSPASVDEAPQPSLPPEAAQSGDLHLW	4740
Qy	4741	EALARKRMSREPTLDSISELPEEDGRSQRLPQEAEEVAPDLSEGYSTADELARTGDADL	4800
Db	4741	EALARKRMSREPTLDSISELPEEDGRSQRLPQEAEEVAPDLSEGYSTADELARTGDADL	4800
Qy	4801	SHTSSDDESRAGTSLVTYILKAGRPQGTSLASKVGAPAAPSVKPOQOQBPLAARPPPLG	4860
Db	4801	SHTSSDDESRAGTSLVTYILKAGRPQGTSLASKVGAPAAPSVKPOQOQBPLAARPPPLG	4860
Qy	4861	DLSTKDLGDPMSDKAAVKIOAAFGYKVRKEMQOEGPMFSHTFGDTEAQVGDALEECV	4920
Db	4861	DLSTKDLGDPMSDKAAVKIOAAFGYKVRKEMQOEGPMFSHTFGDTEAQVGDALEECV	4920
Qy	4921	VASKADVRARWLKDGVELTDGRHHIIDLQGDGTCSLLIAGLDRADAGCYTCQVSNKFGQV	4980
Db	4921	VASKADVRARWLKDGVELTDGRHHIIDLQGDGTCSLLIAGLDRADAGCYTCQVSNKFGQV	4980
Qy	4981	THSACVVVSGSESAESSSGGELDDAFRRARLRHLFRTKSPAESVDSDELFLSADGPA	5040
Db	4981	THSACVVVSGSESAESSSGGELDDAFRRARLRHLFRTKSPAESVDSDELFLSADGPA	5040
Qy	5041	EPEPADWQYRDEDEHFCIRFPALEARQAVTRFQEMFATLIGIGVEIKLVEQGRVEM	5100
Db	5041	EPEPADWQYRDEDEHFCIRFPALEARQAVTRFQEMFATLIGIGVEIKLVEQGRVEM	5100
Qy	5101	CISKETPAPVVPPEPLPSLLTSDAAPVFLTELQNEVQDQGPVPSFDCVWTGQPMPSVRWF	5160
Db	5101	CISKETPAPVVPPEPLPSLLTSDAAPVFLTELQNEVQDQGPVPSFDCVWTGQPMPSVRWF	5160
Qy	5161	KQGLLEEDHYMINEDQQGGHQLIITAVVPADMGVYRCLAEENSMGVSSTKAELRVDLTS	5220
Db	5161	KQGLLEEDHYMINEDQQGGHQLIITAVVPADMGVYRCLAEENSMGVSSTKAELRVDLTS	5220
Qy	5221	TDVDTADATYESSYSAQGYLSRQEGTSTTDEGQLPOVVEELRDLQVAPGTFLAKF	5280
Db	5221	TDVDTADATYESSYSAQGYLSRQEGTSTTDEGQLPOVVEELRDLQVAPGTFLAKF	5280
Qy	5281	QLKVKGYPARLYWFKDQGPLTASAHIRMTGKILHTLEIISVTRSDSGQYAAIISNANG	5340
Db	5281	QLKVKGYPARLYWFKDQGPLTASAHIRMTGKILHTLEIISVTRSDSGQYAAIISNANG	5340
Qy	5341	AAVSSARLLVRGPDEPEEKPADSVHQBQLVPPRMLERFTPKKVKKGSSTITSVKVEGRPV	5400

Db	5341	AAYSABLLVRGDEPEEKPADSVHBQLVPPRMLERFTPKVKVKGSSITTSFKVKEGRPVP	5400
Qy	5401	TVHWLREAEERGVLWTGPTDPTGYTVASSAQOHSVLVLDVGRHOHGTYTTCIASNAAGQALC	5460
Db	5401	TVHWLREAEERGVLWTGPTDPTGYTVASSAQOHSVLVLDVGRHOHGTYTTCIASNAAGQALC	5460
Qy	5461	SASLHVSGLPKVBEOQKVKEALISTFLOQTTOAISAOGLETFASFADLGGORKEEPLAAKE	5520
Db	5461	SASLHVSGLPKVBEOQKVKEALISTFLOQTTOAISAOGFQTASFADLGGORKEEPLAAKE	5520
Qy	5521	ALGHLSLAEVGTTEFPLOKLSQITTEMVSAKITQAKLOVPGGSDSDSKTPSPASPRHGRSR	5580
Db	5521	ALGHLSLAEVGTTEFPLOKLSQITTEMVSAKITQAKLOVPGGSDSDSKTPSPASPRHGRSR	5580
Qy	5581	PSSSIQESSSESDEGDGARGEIPDIYVVTADYILPLGAEQDAITLREGQYVEVLDAAHPLRW	5640
Db	5581	PSSSIQESSSESDEGDGARGEIPDIYVVTADYILPLGAEQDAITLREGQYVEVLDAAHPLRW	5640
Qy	5641	LVRTKTKSPSRQGWVSPAYLDRRLKLSPEWGAAEAPFPGEAUSEDEYKARLSVIOE	5700
Db	5641	LVRTKTKSPSRQGWVSPAYLDRRLKLSPEWGAAEAPFPGEAUSEDEYKARLSVIOE	5700
Qy	5701	LLSSEQAFVEELOFLOSHHLOHLERCPHPVPIAVAGQKAVIPRNRDIGHFHSFLOQLQ	5760
Db	5701	LLSSEQAFVEELOFLOSHHLOHLERCPHPVPIAVAGQKAVIPRNRDIGHFHSFLOQLQ	5760
Qy	5761	CDTDDDVAMCFIKNOAASFQYLEFLVGRVOAESVVVSTAIQBFYKKYAEALLAGDPSP	5820
Db	5761	CDTDDDVAMCFIKNOAASFQYLEFLVGRVOAESVVVSTAIQBFYKKYAEALLAGDPSP	5820
Qy	5821	PPPLQHYLEQPVRYQALIKELIRNKARNRQNCALLEQAQYAVVSALPORAENKLHV	5880
Db	5821	PPPLQHYLEQPVRYQALIKELIRNKARNRQNCALLEQAQYAVVSALPORAENKLHV	5880
Qy	5881	SLMENYPTGLEALGEPIROGHIIVWEGAPGAPMPKGNHRHVLFRNHLVICKPRRDSRT	5940
Db	5881	SLMENYPTGLEALGEPIROGHIIVWEGAPGAPMPKGNHRHVLFRNHLVICKPRRDSRT	5940
Qy	5941	DTVSVYFRNMKLSSIDLNDQVEGDDRAPEVMQEREDSVRKYLLOARTAIKSSWVKETC	6000
Db	5941	DTVSVYFRNMKLSSIDLNDQVEGDDRAPEVMQEREDSVRKYLLOARTAIKSSWVKETC	6000
Qy	6001	GIQORLALPVWRPPDPEEBELADCTAELGETVKLCARVTGTPKPVISWYKDGKAVQVDPHH	6060
Db	6001	GIQORLALPVWRPPDPEEBELADCTAELGETVKLCARVTGTPKPVISWYKDGKAVQVDPHH	6060
Qy	6061	ILIEDPDGSCALILDSLTQVDSQWYCFAPASAGNCSTLGKILVQVPRFRVFNKVRASPVV	6120
Db	6061	ILIEDPDGSCALILDSLTQVDSQWYCFAPASAGNCSTLGKILVQVPRFRVFNKVLASPFV	6120
Qy	6121	EGEDAQFTCTIEGAPVQIRWYKDGALLTTGNKFOTLSEPRSGLLAVITRAAKEDLGLY	6180
Db	6121	EGEDAQFTCTIEGAPVQIRWYKDGALLTTGNKFOTLSEPRSGLLAVITRAAKEDLGLY	6180
Qy	6181	ECELVNRLGSARASAEALRIQSPMLQAOEQCHREQLVAAVEDTTLSR-----ADQEVY	6232
Db	6181	ECELVNRLGSARASAEALRIQSPMLQAOEQCHREQLVAAVEVTEQTKVPKKTIVIEETIT	6240
Qy	6233	SVLKRLLGPKAPGPSTGDTJTGPGCPRG-----APAL-----QETGSP-	6271
Db	6241	TVTKSPRGQRR--SPSKSPSRSPSCSASPLRPGILLAPDLLYLPAGQPRRPEAEPQKPV	6299
Qy	6272	-PVTCGTSEAPVPPRV-----POPLLHEGPE-----QBPPEATARAQEWTVPIRMEGA	6317
Db	6300	VPTLYVTEAHSHPALPGLSGPQKWKVEETIEVRVKMGMQGVSPTE--VPRSSSGH	6357
Qy	6318	AW--PGAGTG 6325	
Db	6358	LFTLPGATPG 6367	
RESULT 2			
Q8NHN3			


```
10 Q8NH3 PRELIMINARY; PRT; 2212 AA.
AC Q8NH3;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Obscurin (Fragment).
GN Name=OBSCN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Young P., Ehler E., Gautel M.;
RL Submitted (May-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Gautel M.S.;
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ314905; CAC85752.1; -.
DR HSP; Q9UQH9; 1E2.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00047; Ig_2.
DR SMART; SM00408; IGC2; 10.
DR PROSITE; PS50835; IG_LIKE; 22.
FT NON TER 1
FT NON TER 2212
SQ SEQUENCE 2212 AA; 241979 MW; B6DFA263EB749465 CRC64;

Query Match 20.6%; Score 8486.5; DB 2; Length 2212;
Best Local Similarity 76.5%; Pred. No. 3.2e-265;
Matches 1692; Conservative 1; Mismatches 2; Indels 517; Gaps 3;

Qy 2734 LHVGITKRLKTMVELEGSCSFECVLSHESADPAMWTGKTVGSSSRFOATROGRKI 2793
Db 1 LHVGITKRLKTMVELEGSCSFECVLSHESADPAMWTGKTVGSSSRFOATROGRKI 60

Qy 2794 LVVREAPSDAGEVVFVSGTSMAMLVIRGASLKDAGEYTCVEASKSLHVEE 2853
Db 61 LVVREAPSDAGEVVFVSGTSMAMLVIRGASLKDAGEYTCVEASKSLHVEE 120

Qy 2854 GTPVHMLKDRKAIRKSKQYDVVCGTSMAMLVIRGASLKDAGEYTCVEASKSLHVEE 2913
Db 121 GTPVHMLKDRKAIRKSKQYDVVCGTSMAMLVIRGASLKDAGEYTCVEASKSLHVEE 180

Qy 2914 KANCFTEELTNLOVEEKGTAFTCKTEHPAATVTRKGLLELRASGKHQPSQEGTLRLT 2973
Db 181 KANCFTEELTNLOVEEKGTAFTCKTEHPAATVTRKGLLELRASGKHQPSQEGTLRLT 240

Qy 2974 ISALEKADSDTYTCDIGQAQRAQLLVQ-GRRVHIIEDLVVDVQEGSSATFRCRISPAN 3032
Db 241 ISALEKADSDTYTCDIGQAQRAQLLVQAGRRVHIIEDLVVDVQEGSSATFRCRISPAN 300

Qy 3033 YEPVHWFIDKTPHANELNEIDAQPGGYHVLTLRQALKDSGTIYFEAGDQASALRV 3092
Db 301 YEPVHWFIDKTPHANELNEIDAQPGGYHVLTLRQALKDSGTIYFEAGDQASALRV 360

Qy 3093 EKPSVFSRELTDAITTEGEDTLVCETSTCDIPMCWTGDKTGRSARCQLSHEGHAQL 3152
Db 361 EKPSVFSRELTDAITTEGEDTLVCETSTCDIPVCWTGDKTGRSARCQLSHEGHAQL 420

Qy 3153 LITGATLQDSGRYKCEAGGACSSSIVRVHARVPFQEAALKDLEVLGGAAATLRCVLSVA 3212
Db 421 LITGATLQDSGRYKCEAGGACSSSIVRVHARVPFQEAALKDLEVLGGAAATLRCVLSVA 480

Qy 3213 APVKWCYGNVLRPGDKYSLRQEGAMLELVVRNLRPQDSGRYSCSFGQDTTSATLVTAL 3272
Db 481 APVKWCYGNVLRPGDKYSLRQEGAMLELVVRNLRPQDSGRYSCSFGQDTTSATLVTAL 540

Qy 3273 PAQFIGKLRNKEATEGATATLRCELSTKPAVPEWRKSGSETLRDGRYCLURQDGMCELQIR 3332
Db 541 PAQFIGKLRNKEATEGATATLRCELSTKPAVPEWRKSGSETLRDGRYCLURQDGMCELQIR 600
```

```
Qy 3333 GLAVWDAAEYSCVCGBERTSASLTIRPMPAHFIGLRHQESIEGATATLRCELSCAAAPVE 3392
Db 601 GLAVWDAAEYSCVCGBERTSASLTIRPMPAHFIGLRHQESIEGATATLRCELSCAAAPVE 660

Qy 3393 WRKGRESLRDGRHSLRQDGAVALCELOICGLAVADAGEYSCVCGBERTSATLTVKALPAKF 3452
Db 661 WRKGRESLRDGRHSLRQDGAVALCELOICGLAVADAGEYSCVCGBERTSATLTVKALPAKF 720

Qy 3453 TEGLRNEEAVEGATAMLCWCELSKVAPVPEWRKGPENLRDGRYILRQEGTRCELQICGLAM 3512
Db 721 TEGLRNEEAVEGATAMLCWCELSKVAPVPEWRKGPENLRDGRYILRQEGTRCELQICGLAM 780

Qy 3513 ADAGEYLCVCGBERTSATLTIRALPARFIEDVKQAEAREGATAVLQCELSNAAAPVPEWRKG 3572
Db 781 ADAGEYLCVCGBERTSATLTIRALPARFIEDVKQAEAREGATAVLQCELSNAAAPVPEWRKG 840

Qy 3573 SETLRDGRYSLRODGTCKELQIRGLAMADTGEYSCVCGBERTSAMLTVRALPIKFTBGL 3632
Db 841 SETLRDGRYSLRODGTCKELQIRGLAMADTGEYSCVCGBERTSAMLTVRALPIKFTBGL 900

Qy 3633 RNEEATEGATAVLRCELSCWAPVPEWRKSGHETLRDGRHSLRQDGAVALCELOICGLAM 3692
Db 901 RNEEATEGATAVLRCELSCWAPVPEWRKSGHETLRDGRHSLRQDGAVALCELOICGLAM 960

Qy 3693 EYLCWCKERTSAMLTVRAMPKFI EYGLRNEEATEGDTATLWCELSKAAAPVPEWRKGHETL 3752
Db 961 EYLCWCKERTSAMLTVRAMPKFI EYGLRNEEATEGDTATLWCELSKAAAPVPEWRKGHETL 1020

Qy 3753 RDGRHSLRQDGRCELQIRGLAVVDAGEYSCVCGBERTSATLTVRALPARFIEDVKQAE 3812
Db 1021 RDGRHSLRQDGRCELQIRGLAVVDAGEYSCVCGBERTSATLTVRALPARFIEDVKQAE 1080

Qy 3813 AREGATAVLQCELSKAAAPVPEWRKSGSETLRDGRYSLRQDGTGRCELQIHGLSVADTGEYSC 3872
Db 1081 AREGATAVLQCELSKAAAPVPEWRKSGSETLRDGRYSLRQDGTGRCELQIHGLSVADTGEYSC 1140

Qy 3873 VCGBERTSATLTVR----- 3886
Db 1141 VCGBERTSATLTVRALPARFTQDLKTKAESEGATATLQCELSKVAPVPEWRKGPETLRDGG 1200

Qy 3887 ----- 3886
Db 1201 RYSLKQDGTGRCELQIHDLSDVADAGEYSCMCGBERTSATLTVRALPARFTEGLRNEEAMEG 1260

Qy 3887 ----- 3886
Db 1261 ATATLQCELSKAAAPVPEWRKSGLEALRDGDGKYSLRQDGAVALCELOIHGLAMADNGVYSCVCGQ 1320

Qy 3887 ----- 3886
Db 1321 ERTSATLTVRALPARFIEDMRNQKATEGATVTLQCKLRKAAPVPEWRKGPNTLKDGRYSL 1380

Qy 3887 ----- 3886
Db 1381 KQDGTSCELQIRGLVIADAGEYSCI CQBERTSATLTVRALPARFIEDVRNHEATEGATAV 1440

Qy 3887 ----- 3886
Db 1441 LQCELSKAAAPVPEWRKSGSETLRDGRYSLRQDGTGRCELQIRGLAVEDTGEYLCVCGBERTS 1500

Qy 3887 ----- 3886
Db 1501 ATLTVRALPARFIDMTNQAEAREGATATLHCELSKVAPVPEWRKGPETLRDGRHSLRODG 1560

Qy 3887 ----- 3886
Db 1561 TRCELQIRGLSVADAGEYSCVCGBERTSATLTIREATEGATAMLCWCELSKVAPVPEWRKGP 1620

Qy 3887 -----APQPVPREPLQ 3897
Db 1621 ETLRDGRYSLRQDGTGRCELQIHGLSVADTGEYSCVCGBERTSATLTVKARQPVPREPLQ 1680
```


QY 3898 SLOAEGSTATLQCELSSEPTATVVMKSGGLQLQANGRRPRELQCGTAELVLQDLOREDTG 3957
Db 1681 SLOAEGSTATLQCELSSEPTATVVMKSGGLQLQANGRRPRELQCGTAELVLQDLOREDTG 1740
QY 3958 EYTCCTGSOATATLTVTAAPVRFLELQHQVEDGGTAHLCCELSRAGASVEWVKGSILQ 4017
Db 1741 EYTCCTGSOATATLTVTAAPVRFLELQHQVEDGGTAHLCCELSRAGASVEWVKGSILQ 1800
QY 4018 LPPCAKYQWQDGAALVVRGVEQEDAGDNYTCDTGHTQSMASLSVRPRPKFKTRLQSL 4077
Db 1801 LPPCAKYQWQDGAALVVRGVEQEDAGDNYTCDTGHTQSMASLSVRPRPKFKTRLQSL 1860
QY 4078 EQETGDIARLCCOLDAESGAVVQMLKEGVELHAGPKYEMRSQGATRELLIHQLEAKDTG 4137
Db 1861 EQETGDIARLCCOLDAESGAVVQMLKEGVELHAGPKYEMRSQGATRELLIHQLEAKDTG 1920
QY 4138 EYACVTGGOKTAASLRVT-EPVTTIVRGLVDAEVTADVEDVFCVSRAGATGVQWCLQG 4196
Db 1921 EYACVTGGOKTAASLRVTGEPEVTTIVRGLVDAEVTADVEDVFCVSRAGATGVQWCLQG 1980
QY 4197 LPQSNVEVAVRGRHTLRLKGVTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPL 4256
Db 1981 LPQSNVEVAVRGRHTLRLKGVTPEDAGTVSFHLGNHASSAQLTVRAPEVTILEPL 2040
QY 4257 QVQLSEGQDASFQRLSRASQEARWALGGVLPQANEMNDITVEQGTLLHLTLHKVTLE 4316
Db 2041 QVQLSEGQDASFQRLSRASQEARWALGGVLPQANEMNDITVEQGTLLHLTLHKVTLE 2100
QY 4317 DAGTVSFHVGTCSSEAOQLKVTAKNTVVRGLENVLEALGEGEALFEQOLSQPEVAHTWLLD 4376
Db 2101 DAGTVSFHVGTCSSEAOQLKVTAKNTVVRGLENVLEALGEGEALFEQOLSQPEVAHTWLLD 2160
QY 4377 DEPVRTSENAEVVFENGLRHLLLNLRPODSCRVTFLAGDMVTSFALTVR 4428
Db 2161 DEPVHTSENAEVVFENGLRHLLLNLRPODSCRVTFLAGDMVTSFALTVR 2212

RESULT 3
Q9HCD3
ID Q9HCD3 PRELIMINARY; PRT; 1319 AA.
AC Q9HCD3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DE K1AA1639 protein (Fragment).
GN Name=K1AA1639;
OS Homo sapiens (Human),
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hiroawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT xviii. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro."
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046859; BAB3465.1; -.
DR HSSP; P56276; 1FHG.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0004648; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR011009; Kinase like.
DR InterPro; IPR000719; Prot_kinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IGC2; 1.

DR PROSITE; PS50853; FN3; 1.
DR PROSITE; PS50835; IG LIKE; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 2.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
FT NON TER 1
SQ SEQUENCE 1319 AA; 142048 MW; FBA43AE17204EF48 CRC64;
Query Match 16 7%; Score 6902; DB 2; Length 1319;
Best Local Similarity 99.8%; Pred. No. 2e-214;
Matches 1319; Conservative 0; Mismatches 0; Indels 2; Gaps 1;
QY 6648 AMGVISYLSLTCSPFFAGESDRATLLNVLEGRVSWSPMAAHLSEDAKDFIKATLQAPQ 6707
Db 1 AMGVISYLSLTCSPFFAGESDRATLLNVLEGRVSWSPMAAHLSEDAKDFIKATLQAPQ 60
QY 6708 ARPSAAOCLSHPWFLKMPABEAHFINTKQLKFLARSRWORSIMSKYKSLVMRSIPELL 6767
Db 61 ARPSAAOCLSHPWFLKMPABEAHFINTKQLKFLARSRWORSIMSKYKSLVMRSIPELL 120
QY 6768 RGPDPSPSLGVARHLCRDTGCGSSSSSSSSDNELAPPARAKSLPPSPVTHSPLLHPRGFLR 6827
Db 121 RGPDPSPSLGVARHLCRDTGCGSSSSSSSSDNELAPPARAKSLPPSPVTHSPLLHPRGFLR 180
QY 6828 PSASLPBEAEASERSTAPAPPASPEGAGPPAAQGCVPVHRSVIRSLFYHQAGESPEHGAL 6887
Db 181 PSASLPBEAEASERSTAPAPPASPEGAGPPAAQGCVPVHRSVIRSLFYHQAGESPEHGAL 240
QY 6888 APGSRHRPARRHLLKGGYTAGALPGLREPLMEHRVLEEAAREEQATLLAKAPSFETAL 6947
Db 241 APGSRHRPARRHLLKGGYTAGALPGLREPLMEHRVLEEAAREEQATLLAKAPSFETAL 300
QY 6948 RLPASGTHLAPGHSHSLSHDPSPTPRPSSEACGEAQLPSAPSGCAPIRDMGHPOGSKOL 7007
Db 301 RLPASGTHLAPGHSHSLSHDPSPTPRPSSEACGEAQLPSAPSGCAPIRDMGHPOGSKOL 360
QY 7008 PSTGCHPGTAQPERPSPDPSPWQGPAPPCHPKQGSAPQEGCSPHAPVAPCPGSPGSCCK 7067
Db 361 PSTGCHPGTAQPERPSPDPSPWQGPAPPCHPKQGSAPQEGCSPHAPVAPCPGSPGSCCK 420
QY 7068 EAPLVPSPTLGOPOAPPAPAKASPLDLSKMGPDLSLGRPKGPCSSPGSASQASSQ 7127
Db 421 EAPLVPSPTLGOPOAPPAPAKASPLDLSKMGPDLSLGRPKGPCSSPGSASQASSQ 480
QY 7128 VSSLRVGSSQVGTPEPGLDAEGWTOEAEDLSSTPTLQRPQOATMRKFSLGRGGYAG 7187
Db 481 VSSLRVGSSQVGTPEPGLDAEGWTOEAEDLSSTPTLQRPQOATMRKFSLGRGGYAG 540
QY 7188 VAGYGTAFGCDAGGMLGGQPMWARIAMAVSQSEEEBQEEARAEBSQSEEQEAREASPLP 7247
Db 541 VAGYGTAFGCDAGGMLGGQPMWARIAMAVSQSEEEBQEEARAEBSQSEEQEAREASPLP 600
QY 7248 QVSARPVVGRAPTRSSPEPTWEDIGVSLVQIRDLSDGAADAADTISLDISEVDPAYL 7307
Db 601 QVSARPVVGRAPTRSSPEPTWEDIGVSLVQIRDLSDGAADAADTISLDISEVDPAYL 660
QY 7308 NLSLDYDIKYLPEEFMIFRKVPKSAQPEPPSPMAEEELAEFFPTWPGELGPHAGLEI 7367
Db 661 NLSLDYDIKYLPEEFMIFRKVPKSAQPEPPSPMAEEELAEFFPTWPGELGPHAGLEI 720
QY 7368 TEESDVEDALLAEAAVGRKKWSPRSRSLPHFPGRHLPDDEPAELGLRERVKASVEHISR 7427
Db 721 TEESDVEDALLAEAAVGRKKWSPRSRSLPHFPGRHLPDDEPAELGLRERVKASVEHISR 780
QY 7428 ILKGRPEGLEKEGPPRRKKPGLASFRLSGLKSWDRAPTFRLSDETVVLGQSVTLAQCVS 7487
Db 781 ILKGRPEGLEKEGPPRRKKPGLASFRLSGLKSWDRAPTFRLSDETVVLGQSVTLAQCVS 840
QY 7488 AQPAAQATWSKOGAPLESSSRVLSATLKNFQLLTLLVVAEDLGVYTCVSNALGTVTT 7547
Db 841 AQPAAQATWSKOGAPLESSSRVLSATLKNFQLLTLLVVAEDLGVYTCVSNALGTVTT 900
QY 7548 TGVLRKAERPSSPCPDIGEVADGVLLVWKPVESGPTVYIVQCSLEGGSWTTLASDIF 7607


```
Db 901 TGVLRKAERBSSPCPDIGEVYADGVLLVWKPVSYPVTVIYVQCSLEGGSWTTLASDIP 960
Qy 7608 DCCYLTSLKRGGRYTTTRTACVSKAGMPYSSPSEQVLLGGPSPHSLAEEESQGRSAQPLP 7667
Db 961 DCCYLTSLKRGGRYTTTRTACVSKAGMPYSSPSEQVLLGGPSPHSLAEEESQGRSAQPLP 1020
Qy 7668 STKTFATQTOQRGFSVRQCKEASGRALAAKIIIPVHPKDKTAVLREYALKGLRHPH 7727
Db 1021 STKTFATQTOQRGFSVRQCKEASGRALAAKIIIPVHPKDKTAVLREYALKGLRHPH 1080
Qy 7728 LAQLHAAVYLSPRHLVLIILELCSGPELLPCLAEASYSSESEVKYLDWMLSATQYLHNRQHI 7787
Db 1081 LAQLHAAVYLSPRHLVLIILELCSGPELLPCLAEASYSSESEVKYLDWMLSATQYLHNRQHI 1140
Qy 7788 LHLDRSENMIITYNLLKVVDLNAOGLSSEKVLPSDKFKDYLETWAPLELGGGAVPQ 7847
Db 1141 LHLDRSENMIITYNLLKVVDLNAOGLSSEKVLPSDKFKDYLETW--ELLEGGGAVPQ 1198
Qy 7848 TDIWAIGTAFIMLSAEVPSSESGARDLQRLKGLVRLSRCYAGLSGGGAVAFRLSTLCA 7907
Db 1199 TDIWAIGTAFIMLSAEVPSSESGARDLQRLKGLVRLSRCYAGLSGGGAVAFRLSTLCA 1258
Qy 7908 QPMGRPCASSCLOCPMLTEEGPACSRPAPVTFPTARLRFVVRNREKRALLYKRHNLAQV 7967
Db 1259 QPMGRPCASSCLOCPMLTEEGPACSRPAPVTFPTARLRFVVRNREKRALLYKRHNLAQV 1318
Qy 7968 R 7968
Db 1319 R 1319

RESULT 4
Q9HCL6 ID Q9HCL6 PRELIMINARY; PRT; 1596 AA.
AC Q9HCL6;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE KIAA1556 protein (Fragment).
GN Name=KIAA1556;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hiroseawa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT xviii. The complete sequences of 100 new cDNA clones from brain which
RL DNA Res. 7:273-281(2000).
DR EMBL; AB046776; BAB13382.1; --
DR HSSP; P56276; 1FHG
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00408; IGC2; 7.
DR PROSITE; PS50835; IG_LIKE; 15.
FT NON_TER 1
FT NON_TER 1596
SQ SEQUENCE 1596 AA; 175279 MW; 88937B0760C4E2EF CRC64;

Query Match
Best Local Similarity 81.7%; Score 6619.5; DB 2; Length 1596;
Matches 1309; Conservative 70; Mismatches 203; Indels 21; Gaps 7;

Qy 2726 RARVRVHDLVHGITKRLKTMELVEGESCSFECVLSHESASDPAMWTVGKTVGSSSRFQA 2785
Db 1 RARVRVHDLVHGITKRLKTMELVEGESCSFECVLSHESASDPAMWTVGKTVGSSSRFQA 60
Qy 2786 TRQGRKYLIVREAA PSDAGEVWFSVRGLTSKASLIVRERPAALIKPLEDQWVAPGEDVE 2845
```

```
Db 61 TRQGRKYLIVREAA PSDAGEVWFSVRGLTSKASLIVRERPAALIKPLEDQWVAPGEDVE 120
Qy 2846 LRCELSRAGTPVHMLKDKRKAIKRSQKYDVVCEGTMAMLVIRGASLKDAGEYTCVEASKS 2905
Db 121 LRCELSRAGTPVHMLKDKRKAIKRSQKYDVVCEGTMAMLVIRGASLKDAGEYTCVEASKS 180
Qy 2906 TASHLVEEKANCFTEELTNLQVEKGTAVFTCKTEHPAATVTVWRKGLLELRASGHQPSQ 2965
Db 181 TASHLVEEKANCFTEELTNLQVEKGTAVFTCKTEHPAATVTVWRKGLLELRASGHQPSQ 240
Qy 2966 EGLTLRLTISALEKADSDTYTCDIGQAQSRQAQLLVQGRRVHIIEDLEDVDVQEGSSATFR 3025
Db 241 EGLTLRLTISALEKADSDTYTCDIGQAQSRQAQLLVQGRRVHIIEDLEDVDVQEGSSATFR 300
Qy 3026 CRISPANYEPVHMFDPKTPHANELNEIDAPQPGYHVLTLRQALAKDSGTTFYFAGDORA 3085
Db 301 CRISPANYEPVHMFDPKTPHANELNEIDAPQPGYHVLTLRQALAKDSGTTFYFAGDORA 360
Qy 3086 SAALRVTEKPSVFSRELTDTATITEGEDLTLCVETSTCDIPMCWTGDKGKTLRGSARCOLSH 3145
Db 361 SAALRVTEKPSVFSRELTDTATITEGEDLTLCVETSTCDIPVCWTGDKGKTLRGSARCOLSH 420
Qy 3146 EGHRQAQLLITGATLQDSGRYKCEAGGACSSSIVRVHARPVRFQBALKDLEVLGGAATLR 3205
Db 421 EGHRQAQLLITGATLQDSGRYKCEAGGACSSSIVRVHARPVRFQBALKDLEVLGGAATLR 480
Qy 3206 CVLSSVAAAPVKWCYGNVLRPGDKYLSRQSGAMLELVVRNLRPQDSGRYSCSFGDQTSTA 3265
Db 481 CVLSSVAAAPVKWCYGNVLRPGDKYLSRQSGAMLELVVRNLRPQDSGRYSCSFGDQTSTA 540
Qy 3266 TLTVTALPAQFIGKLRNKEATEGATATLRCELSKTAPVWRKGSSETLRDGRYCLRDGA 3325
Db 541 TLTVTALPAQFIGKLRNKEATEGATATLRCELSKAAAPVWRKGSSETLRDGRYCLRDGA 600
Qy 3326 MCELOIRGLAMVDAAYSCVCGEERTSASLTIRPMPAHFIFGLRHQESIEGATATLRCEL 3385
Db 601 MCELOIRGLAMVDAAYSCVCGEERTSASLTIRPMPAHFIFGLRHQESIEGATATLRCEL 660
Qy 3386 SKAAPVWRKGRESLRDGRHSLRQDGAVALCELOICGLAVADAGEYSCVCGEERTSATLTV 3445
Db 661 SKAAPVWRKGRESLRDGRHSLRQDGAVALCELOICGLAVADAGEYSCVCGEERTSATLTV 720
Qy 3446 KALPAKTEGLRNEEAEGGATAMLCESKVAPVWRKGPENLRDGRYILRQEGTRCEL 3505
Db 721 KALPAKTEGLRNEEAEGGATAMLCESKVAPVWRKGPENLRDGRYILRQEGTRCEL 780
Qy 3506 QICGLAMADAGEYLCVCGQERTSATLTIRALPARFIEDVKNOEAREGATAVLQCELNSAA 3565
Db 781 QICGLAMADAGEYLCVCGQERTSATLTIRALPARFIEDVKNOEAREGATAVLQCELNSAA 840
Qy 3566 PVWRKGSSETLRDGRYSLRQDGTCKELQIRGLAMADTGEYSVCVCGERTSAMLTVRALP 3625
Db 841 PVWRKGSSETLRDGRYSLRQDGTCKELQIRGLAMADTGEYSVCVCGERTSAMLTVRALP 900
Qy 3626 IKPTEGLRNEEAEGGATAMLCESKVAPVWRKGPENLRDGRHSLRQDGRARCELQIRG 3685
Db 901 IKPTEGLRNEEAEGGATAMLCESKVAPVWRKGPENLRDGRHSLRQDGRARCELQIRG 960
Qy 3686 LVAEDAGEYLCMCKERTSAMLTVRAMPSPKFIEGLRNEEAEGGATATLWCELSKAAAPVW 3745
Db 961 LVAEDAGEYLCMCKERTSAMLTVRAMPSPKFIEGLRNEEAEGGATATLWCELSKAAAPVW 1020
Qy 3746 RKGHETLRDGRHSLRQDGRARCELQIRGLAVVDAGEYSCVCGERTSATLTVRALPARFI 3805
Db 1021 RKGHETLRDGRHSLRQDGRARCELQIRGLAVVDAGEYSCVCGERTSATLTVRALPARFI 1080
Qy 3806 EDVKNOEAREGATAVLQCELNSKAAAPVWRKGSSETLRDGRYSLRQDGTCKELQIRGLSVA 3865
Db 1081 EDVKNOEAREGATAVLQCELNSKAAAPVWRKGSSETLRDGRYSLRQDGTCKELQIRGLSVA 1140
Qy 3866 DTGEYSVCVCGERTSATLTVRAPQPVFPREPLQSLQAEBSGSTATLQCELSSEPTATVWSKG 3925
Db 1141 DTGEYSVCVCGERTSATLTVRALPARFTQDLKTBESGATATLQCELSK-VAPVWKKG 1199
```


Db 813 GYGTFAGDAGGMLGQGLPWARMAVWSQS-----SEEQDEAATESPQPLE 859

Qy 7250 SARVPVEVGRAPTRSSPEPTWEDIGQVSLVOIRDLSDGAEAAADTISLDISEVPAYLNL 7309

Db 860 SLGPIAASGVPLRTSPSLTPEWEVEQVSLVQIRDLSDGAEAAADTISLDISEVPAYLNL 919

Qy 7310 SLDYDIKYLPEEFMIFERKVPKSAQPEPPSPMAESEE---LAEP-PEPTWPPWPGELGPHAG 7364

Db 920 SLDYDIKYLPEEFMIFERKVPKSAQPEPPSPMAESEE---LAEP-PEPTWPPWPGELGPHAG 978

Qy 7365 LEITEESE---DVALAALAAVGRKRKWSRSRSLFHPGGRHLPLVVAEDLGVYTCVSNA 7421

Db 979 LEITEEPEDGLEALLGANAVGRKRW-SFSGRLFPQPGRCLSGSEEPVELGLQRVKAS 1037

Qy 7422 VEHLISRLKRGPEGLKEGPRKXPLGLASFRSLGLKSWDRAPTEFLRELSDETIVLQGSVT 7481

Db 1038 MAHISRLKRGPEGEREPGRKXAGLASFRLSLGLKGRD-----QELSDEAVVLQGSVT 1091

Qy 7482 LACOVSAQPAQAATWSKDGAPLESSRVLISATLKNFOLLTILVVAEDLGVYTCVSNA 7541

Db 1092 LACQVLAQPTAQAATWSKDGVLLESSGHLLISSTLKNFOLLTILVVAEDLGVYTCVSNA 1151

Qy 7542 LGTIVTTGVLKARPPSSPCPDIGEVYADGVLLVWKPEVSGPVYIVQCSLEGGSWTT 7601

Db 1152 LGTAVTTGVLKARPPSSPCPDIGEVYADGVLLVWKPEVSGPVYIVQCSLEGGSWTT 1211

Qy 7602 LASDIFDCYLTSLSRGTYTFTACVSKAGMGYSPPSQVLLGGPSHLASBEE-SQG 7660

Db 1212 LASDISDCYLTGKLSRGMTIFRTACVSKAGMGYSPPSQVLLGGPSHLASBEE-SRG 1271

Qy 7661 RSAQPLSTKTFATQTOIORGRFSVVRQCEKASGRALAAKII PVHPKDKTAVLREYAL 7720

Db 1272 RPAQLLSTKTFATQTOIORGRFSVVRQCEKASGRALAAKII PVHPKDKTAVLREYAL 1331

Qy 7721 KGLRHPHQAOLHAAYLSRHLVLELCSGPELLPCLAEASYSSESEVKDYLMQWLSATQ 7780

Db 1332 KRLHHPHQAOLHAAYLSRHLVLELCSGPELLPCLAEASYSSESEVKDYLMQWLSATQ 1391

Qy 7781 YLHNOHLHLDLRSENMIITEYNLLKVVDLGNAGSLQSEKVLPSDKFKDYLETWAPELLE 7840

Db 1392 YLHAQHLHLDLRSENMMVTEYNLLKVVDLGNAGSLQSEKVPAPENFKDYLETWAPELLE 1451

Qy 7841 GQAVPQTDIWAIGVTAFLMSAEVPSSEGRADLQRLGRLVLSRCYAGLSGAVAF 7900

Db 1452 GQAVPQTDIWAIGVTAFLMSAEVPSSEGRADLQRLGRLVLSRCYAGLSGAVAF 1511

Qy 7901 LRSTLCAQWGRPCASSCLQCPWLTTEGPACSRPAPVTFPTARLRFVVRNREKRRALLYK 7960

Db 1512 LQSSICAQWGRPCASTCLQCGWLTTEGPTGSRPTPTVTFPTVRLRAFVREKRRALLYK 1571

Qy 7961 RHNLAQVR 7968

Db 1572 KHNLAQVR 1579

RESULT 6

Q8NH2 PRELIMINARY; PRT; 1040 AA.

AC Q8NH2; 01-OCT-2002 (Tremblrel. 22, Created)

DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)

DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)

DE Obscurin (Fragment)

GN Name=OBSCN;

OS Homo sapiens (Human)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RA Young P., Ehler E., Gautel M.;

RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Gautel M.S.;

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AJ314906; CAC85753.1; -

DR HSSP; P56276; 1FHG.

DR InterPro; IPR003962; FNIII subd.

DR InterPro; IPR003961; FN III

DR InterPro; IPR008957; FN_III-like.

DR InterPro; IPR007110; IG_c2.

DR InterPro; IPR003598; IG_c2.

DR InterPro; IPR000048; IQ_region.

DR Pfam; PF00041; fn3; 1.

DR Pfam; PF00047; ig; 1.

DR Pfam; PF00612; iq; 1.

DR PRINTS; PR00014; FNTYPEIII.

DR SMART; SM00060; FN3; 1.

DR SMART; SM00408; IGC2; 3.

DR SMART; SM00015; IQ; 1.

DR PROSITE; PS50853; FN3; 1.

DR PROSITE; PS50835; IG_LIKE; 5.

DR PROSITE; PS50096; IQ; 1.

FT NON_TER 1

FT NON_TER 1040 1040

SQ SEQUENCE 1040 AA; 3DB879CA266D7F44 CRC64;

Query Match 13.0%; Score 5358.5; DB 2; Length 1040;

Best Local Similarity 99.4%; Pred. No. 8.1e-165;

Matches 1034; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

Qy 4430 WRLEILPLKNAAYRAGAQAARFTCTLSEAVPVGASWYINGAAVQDPDSDMTVTADGSHQ 4489

Db 1 WRLEILPLKNAAYRAGAQAARFTCTLSEAVPVGASWYINGAAVQDPDSDMTVTADGSHH 60

Qy 4490 ALLRSQAQPHHAGVTTACRDASARLTVLGLPDPPEDEAVHSHSTVTLSWAAPMSD 4549

Db 61 ALLRSQAQPHHAGVTTACRDASARLTVLGLPDPPEDEAVHSHSTVTLSWAAPMSD 120

Qy 4550 GGGGLCGYRVEVKEGATGQWRLCHELVPGPCVVDGLAPGETYFRVAAGVPGVAGEPVH 4609

Db 121 GGGGLCGYRVEVKEGATGQWRLCHELVPGPCVVDGLAPGETYFRVAAGVPGVAGEPVH 180

Qy 4610 LPQTVRLAEPKPKVPPOPSAPESQVAAGEDVSLSEVVAEAGVVIHKGHERIQPGGRF 4669

Db 181 LPQTVRLAEPKPKVPPOPSAPESQVAAGEDVSLSEVVAEAGVVIHKGHERIQPGGRF 240

Qy 4670 EVWSQGRQQLMVIKGTAEQDQGEYHCGLAQSGICPAAATFQVALSPASVDEAPQSLPPE 4729

Db 241 EVWSQGRQQLMVIKGTAEQDQGEYHCGLAQSGICPAAATFQVALSPASVDEAPQSLPPE 300

Qy 4730 AAQEGDLHLWEALARKRMSREPTLDSISELPEEDGRSQRLPQAEAEVAPDLSEGYSTA 4789

Db 301 AAQEGDLHLWEALARKRMSREPTLDSISELPEEDGRSQRLPQAEAEVAPDLSEGYSTA 360

Qy 4790 DELARTGDADLSHTSSDDESAGTPSLVTVLKKAGRPCTSPASKVGAAPAPSVKPPQOOQ 4849

Db 361 DELARTGDADLSHTSSDDESAGTPSLVTVLKKAGRPCTSPASKVGAAPAPSVKPPQOOQ 420

Qy 4850 EPLAAVRPPLGDLSTKDLGDFSMDKAAVKIQAAFKYKVRKEMKQOEGPMFSHTFGDTEA 4909

Db 421 EPLAAVRPPLGDLSTKDLGDFSMDKAAVKIQAAFKYKVRKEMKQOEGPMFSHTFGDTEA 480

Qy 4910 QVGDALRLCECVVASKADVRAARWLKDGVELTIDGRHHHIDQLGDGTCSLLIAGLDRADAGCY 4969

Db 481 QVGDALRLCECVVASKADVRAARWLKDGVELTIDGRHHHIDQLGDGTCSLLIAGLDRADAGCY 540

Qy 4970 TCQVSNKFGVTHSACVVVSGSESESSSGELDDAFRRARLRHLFRKTSKSPAEVSDE 5029

Db 541 TCQVSNKFGVTHSACVVVSGSESESSSGELDDAFRRARLRHLFRKTSKSPAEVSDE 600

Qy 5030 ELFLSADGPAEPPEPADWQTYRDEHFICIRFEALTEARQAVTRFQEMFATLIGIVBIK 5089

Db 601 ELFLSADGPAEPPEPADWQTYRDEHFICIRFEALTEARQAVTRFQEMFATLIGIVBIK 660


```
QY 5090 LIVEQPRRVMCISKETPAPVPPBPLPSLLTSDAAPVFLTELQNEQVODGYVPSFDCVW 5149
DB 661 LIVEQPRRVMCISKETPAPVPPBPLPSLLTSDAAPVFLTELQNEQVODGYVPSFDCVW 720
QY 5150 TQQPMPSVRWFKDGLLEDHDMYLNEDQOGHQLIITAVVPADMGVYRCIAENSMGVSS 5209
DB 721 TQQPMPSVRWFKDGLLEDHDMYLNEDQOGHQLIITAVVPADMGVYRCIAENSMGVSS 780
QY 5210 TKAELRVDLTSDYDTAADAATESSSYFSAQGYLS-SREQEGTESTTDEGQLPQVVEELRD 5268
DB 781 TKAELRVDLTSDYDTAADAATESSSYFSAQGYLS-SREQEGTESTTDEGQLPQVVEELRD 840
QY 5269 LQVAPGTRLAKPQLKVGYPAPRLWFKDGOPLTASAHIRMTGKILHTLIIISVTREDS 5328
DB 841 LQVAPGTRLAKPQLKVGYPAPRLWFKDGOPLTASAHIRMTGKILHTLIIISVTREDS 900
QY 5329 GOYAAIYISNAGAAVSSARLLVRGDEPEEKPADSVHEQLVPPRMLEFPTPKYVKKGSSI 5388
DB 901 GOYAAIYISNAGAAVSSARLLVRGDEPEEKPADSVHEQLVPPRMLEFPTPKYVKKGSSI 960
QY 5389 TFSVKVEGRPVPTVHMLREEAERGVLWIGPDPGTGYTVASSAQOHSVLVLLDVGRQHOGTYT 5448
DB 961 TFSVKVEGRPVPTVHMLREEAERGVLWIGPDPGTGYTVASSAQOHSVLVLLDVGRQHOGTYT 1020
QY 5449 CIASNAAGQALCSASLHVSG 5468
DB 1021 CIASNAAGQALCSASLHVSG 1040

RESULT 7
Q86YC7 ID Q86YC7 PRELIMINARY; PRT; 767 AA.
AC Q86YC7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to titin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC043916; AAH43916.1; -.
DR HSSP; P56276; 1FHG.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR Pfam; PF00047; Ig; 1.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS50835; IG_LIKE; 5.
FT NON_TER
SQ SEQUENCE 767 AA; 84732 MW; 39AB316B63964DAA CRC64;

Query Match 9.4%; Score 3865; DB 2; Length 767;
Best Local Similarity 99.5%; Pred. No. 7.4e-117;
Matches 751; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2248 VKVQGRYTYLLYRVLAEDAGEIQFVAENAESRAQLRVKELPVTLVRLDKKIAMEKHG 2307
DB 5 IRQEGRTYTYLLYRVLAEDAGEIQFVAENAESRAQLRVKELPVTLVRLDKKIAMEKHG 64

QY 2308 VLEQVSRASQVRFWFGSQSLQPGKVELVSDGLYRKLIIISDVHAEDEDTYTCADGVK 2367
DB 65 VLEQVSRASQVRFWFGSQSLQPGKVELVSDGLYRKLIIISDVHAEDEDTYTCADGVK 124

QY 2368 TSAQFFVEQISITIVRGLQDVTVMEPAPAWFECETSIPIVPPKWLIGKTVLQAGNVGL 2427
DB 125 TSAQFFVEQISITIVRGLQDVTVMEPAPAWFECETSIPIVPPKWLIGKTVLQAGNVGL 184

QY 2428 EQEGTVHRLMLRRCSTWMTGPHFTVSKRSSARLVVSDIPVLTTRPLEPKTRELQSVV 2487
```

```
DB 185 EQEGTVHRLMLRRCSTWMTGPHFTVSKRSSARLVVSDIPVLTTRPLEPKTRELQSVV 244
QY 2488 LSCDFRPAPKAVQWYKDDTPLSPSEKFKMSLEGQMAELRLRLMPADAGVYRCOAGSAHS 2547
DB 245 LSCDFRPAPKAVQWYKDDTPLSPSEKFKMSLEGQMAELRLRLMPADAGVYRCOAGSAHS 304
QY 2548 STEVTVEAREVTVTGPLQDAEATEEGWASFSCELSHDEDEEVSNGLMPLYNDSEFHEISH 2607
DB 305 STEVTVEAREVTVTGPLQDAEATEEGWASFSCELSHDEDEEVSNGLMPLYNDSEFHEISH 364
QY 2608 KGRHRTLVLKSIQRADAGIVRASLSKVSTSALEVRVKPVVFLKALDDLSAEERGTLLAQ 2667
DB 365 KGRHRTLVLKSIQRADAGIVRASLSKVSTSALEVRVKPVVFLKALDDLSAEERGTLLAQ 424
QY 2668 CEVSDPBAHVVRKQGVOLGSPDKYDFLHTAGTGLVVDVSPEDAGLYTCHVGSEETRA 2727
DB 425 CEVSDPBAHVVRKQGVOLGSPDKYDFLHTAGTGLVVDVSPEDAGLYTCHVGSEETRA 484
QY 2728 RVRVHDLHVGIITKRLKTMEVLEGESCSFEVLSHESASDPAMWTVGGKTVGSSSRFOATR 2787
DB 485 RVRVHDLHVGIITKRLKTMEVLEGESCSFEVLSHESASDPAMWTVGGKTVGSSSRFOATR 544
QY 2788 QGRKYILVREAAAPSADAGEVVFVSRGLTSKASLIIVRERPAALIKPLEDQWVAPGEDVELR 2847
DB 545 QGRKYILVREAAAPSADAGEVVFVSRGLTSKASLIIVRERPAALIKPLEDQWVAPGEDVELR 604
QY 2848 CELSRAGTPVHMLKDKRAIRKSOKYDVVCEGTWMLVIRGASLKDAGEYTCVEASKSTA 2907
DB 605 CELSRAGTPVHMLKDKRAIRKSOKYDVVCEGTWMLVIRGASLKDAGEYTCVEASKSTA 664
QY 2908 SLHVEKANCFTBELTNLQVEEKGTAVFTCKTSHPAATVTRKGLLELRASGKHQPSQEG 2967
DB 665 SLHVEKANCFTBELTNLQVEEKGTAVFTCKTSHPAATVTRKGLLELRASGKHQPSQEG 724
QY 2968 LTURLTISALEKADSDTYTCDIQOASRAQLLVQG 3002
DB 725 LTURLTISALEKADSDTYTCDIQOASRAQLLVQG 759

RESULT 8
Q8WZB3 ID Q8WZB3 PRELIMINARY; PRT; 2692 AA.
AC Q8WZB3;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE N2B-Titin Isoform.
GN Name=TTN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20309627; PubMed=10850961;
RA Freiburg A., Trombitas K., Hell W., Cazorla O., Fougereousse F.,
RA Centner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labeit S.;
RT "Series of exon-skipping events in the elastic spring region of titin
RT as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21573839; PubMed=11717165;
RA Bang M.L., Centner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labeit D., Gregorio C.C., Granzier H.,
RA Labeit S.;
RT "The complete gene sequence of titin, expression of an unusual ~700
RT kDa titin isoform and its interaction with obscurin identify a novel
RT Z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
RN [3]
```


RP SEQUENCE FROM N.A.
RA Centner S.B.,
RL Submitted (Jul-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277892; CAD12455.1; -
DR HSSP; Q10466; 1G1C.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004896; F:hematopoietin/interferon-class (D200-domain...; IEA.
DR GO; GO:0004674; P:protein serine/threonine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR Pfam; PF00041; fn3; 132.
DR Pfam; PF00047; ig; 3.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00080; FN3; 132.
DR SMART; SM00408; IGC2; 23.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_1.
DR PROSITE; PS00853; FN3; 132.
DR PROSITE; PS00835; IG_LIKE; 88.
DR PROSITE; PS00290; IG_MHC; UNKNOWN_1.
DR PROSITE; PS00435; PEROXIDASE_1; UNKNOWN_1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
SQ SEQUENCE 26926 AA; 2993021 MW; 169AB42637A7C1FB CRC64;

Query Match 8.7%; Score 3580.5; DB 2; Length 26926;
Best Local Similarity 20.3%; Pred. No. 8.2e-106;
Matches 1935; Conservative 1287; Mismatches 3655; Indels 2667; Gaps 322;

QY 4 PQFGAPFLTRKPAFVSVGKDATLSQIVGNTPQVSWEKDQOPVTAGARFLA---Q 60
DB 1030 PGEAPAFYITKPVVQKLVGGVFGQVGNPKPHVYWKSGVPLTGTGYRVKSVYKQ 1089
QY 61 DGDLYRLTLDALGDSQYVCRARNATGEAFVAVGL-----QVD 100
DB 1090 TGEK-KLVISMTFADDAGEYITVVRNKHGETSASALLSEADYELLMKSQEMLYQTQVT 1148
QY 101 A---EACABQAPHL----- 113
DB 1149 AFVQEPKVGETAPGFVSYEYKEKEQALIRKMAKDVTVVVRTVVEDEPHISSFERL 1208
QY 114 -----LRPTSIRVREGSEATFRCV 133
DB 1209 IKETRYRIKTTLEELDEEGEKXAVDISSEAVESGFDLRKNYRILEGNGVTFHCKM 1268
QY 134 GGSPPRAVSWKDGRLGEPDGPVRVRELGEASA-LRIRAAARPDGCTGYEVRAENPLGA 192
DB 1369 SGYPLPKIAWKDGKRI--KHGERVQMDFLQDGRASLRIPVVLPEDEGIYTAPASNIKN 1326
QY 193 ASAAALVVDSDAADTASR--PGTSTAALLAHLQRR---REAMRAEGAPASFP----- 240
DB 1327 AICSGKLVVEPAAPLGAPTYITPLEPVSIRLSLSPRSVSRSPIRMSPARMSPARM 1386
QY 241 -----STCTR-----TCTVTEGKHARLSCTYTGEPKPTVWKK 273
DB 1387 SPARKSPRRLEETDESOLERYKPVFLKPVSKLEGQTAEDFLKVVGRPMPTFWFH 1446
QY 274 DGLVLTEGRRH--VYVEDAQENFVLKILFCQSDRGLYTCTASNLVGOTYSVLVV--- 328
DB 1447 DGOQIVNDYTHKVIKEDGTQSLI--IVPATPSDSEWTVAQNAGRASSISVILTVEAV 1504
QY 329 ----- 328
DB 1505 EHQVKPMFEVEXLKNVNIKEGSOLEKMKVRATGNPNPDIIVLKNSDIIVPHKYPKIRIEGTK 1564
QY 329 -----R 329
DB 1565 GEALKIDS TVSQSAWYTATINAKGRDTRCKVNVVEFAEPEPERKLIIPRGTYRAK 1624
QY 330 EPAPV-----FKKRLQDLEVRKESATFLCE---VP 357
DB 1625 EIAAPELEPLHLRYGOEQWEGDLYDKEQKQKPPFKKLTSLRKRFQPAHFECLTPIG 1684

QY 358 QPSTEAAWFKETRL-----WASAKYGI-----BEEGTERRLT 390
DB 1685 DPTWVVEWLDHGKPLEAANRLRMINEGYCLDYGAVSRDSGIITCRATNKYGTHTSA 1744
QY 391 VRNVSADDAVYICETPEGSR-----TVASLAVQGNL-----LRKLPRKT 430
DB 1745 TLIIVDEKSLVESQLPEGRKGLQRIEELERMAHEGALTGVTTQKEKQKDPDILVPEPV 1804
QY 431 AVRVDGTAMP-CVELAVPVGPVHWRNQEEVVAGRVASISAEGRTHLTITISOCCLVEDVQ 489
DB 1805 RVLEGETARFCRVGTGYPQPKVNNYLNQLIRKSRFRVYDGI-HYLDIVDCKSYDTGE 1863
QY 490 VAFVAG-----DCQTSFRFCVSAARKPLQ-----PVD 518
DB 1864 VKVTAENPEGVIEHKVXLEIQOREDFRSLVRRAPEPRPEFHVHBEFGKLQFEVQKVRD 1923
QY 519 -----PVVKARMESSVILSNPPHGE-----RPVITDGY-----LVKKKXKLGTWTW 560
DB 1924 TTETKEVVKLKRAERITHEKVPESSELRSEKFKRTEEGYVEAITAVELKSRKXDESVEE 1983
QY 561 I---RCHAEAWATPELTVAD---VABEGNFQFRVSALNSFGQSPYLEFPFGTVHLAPKL 613
DB 1984 LLRTKDELLHW--TKELTEBEKKALAEKGITITPTFKPKIELSPSME-----APKI 2034
QY 614 AVRTPLKAVQAVEGGEVTFSDLTVASAG-----EWFLDQALKASS-VYEHCDRTRH 666
DB 2035 FERIQSQTG--QGSDAHFRVR---VGKPDPECEMYKNGVKIERSDRIYWPEDNVC 2088
QY 667 TLTIREVPASLHGQALKF---VANGTESSIRMEVRAAPGLTANKPPAAAAAREVLARLHE 722
DB 2089 ELVIRDTVAE--DSASIMWKAINIAGETSSHAFLVQAKLITFTQ-----ELQDVVAKED 2143
QY 723 E-AQLLAELSQAAAVTWLKDGRTLSPGPKYEQASAGRRVLLVRDVARDDAGLYECV-- 779
DB 2144 TMAITFECETSEPPFKVKWKYKDMGEVHEGDKYRMHSDRKVHFLSLITITDSDAEDYSCVLV 2203
QY 780 -SRGRTAYQLSVQGLARPLHKDMAGSCVDVAGGPAQFECETSEAHVHVHWYKDMELG 838
DB 2204 EDENVKTTAKLIVGAVVVFVKELQD--IEVPESYSGELECIVSPENIEGKWHYNDVELK 2261
QY 839 HSGRFRLEQEDVGTGRHLVAAVTVTQDEGTYSRVEDSDVDFRLRVSEPKVVFVAKELARR 898
DB 2262 SNGKYITTSRG-RQNLTVKDVTKEDQGEYSFVTDGKTTCKLKM-KRPPAILQGLSDQ 2319
QY 899 KLOAEAGASATLSCVAAQATEVTWYKDGKLSSSSKYCMBEATCTRLVAVQQAQADAG 958
DB 2320 KVC--EGDIVQLEVKVLSLESVEGVWMDQGEVQPSDRVHIVIDQSHMLLIETMTKEDAG 2377
QY 959 EYS-----CEAGGQRLSHFLDYKPKVPAQDVAAHSEVQAEAGANATLSCVAAQ- 1010
DB 2378 NYSFTIPALGLSTSGRVSVSDVITPL-----KQYNVIEGTKAVLECKVSVSDV 2427
QY 1011 AEVVMYKDGKLSLSSKLVHVEAKCRRLVVOQAGKTDAGDYSCEARGQVSRFLHITEP 1070
DB 2428 TSVKMYLNDEQIKPDDRVQALVKGTQKORLVNIRTHASDEGYKLVGRVETCNLSVSKI 2487
QY 1071 KMMPAKEQSVHNEVQAEAGASAMLSCEVAAQATEVTWYKDGKLSLSSSKYCMGKVGCTRR 1130
DB 2488 KIIRGLRDLTCTETQ-----NVVFEVELSHSGIDVLMNFKDKKEIKPSSKYIEAHGKIYK 2542
QY 1131 LVLPOAGKADAGEYSCEAGQGVSHLHITPEKGVFAKEQSVHNEVQAEAGTAMLSCEV 1190
DB 2543 LTVLNMKMDDEGKTYTFAGENITSGKL--TVAGGAISK--PLTDQTVABES-QEAVFECEV 2597
QY 1191 AQPOTVETWYKDGKLSLSSSKVRMEVKGCTRLRVVQVGVKADAGEYSCEAGQVSRVSLH 1250
DB 2598 ANPSKGEWLKDGKHLPLTNIRSESGHKKRLIATKLDIDIGYITKVTATSKTSAKUK 2657
QY 1251 ITEPKAVFAKEQLVHNEVTEAGASATLSCVAAQATE-VTWYKDGKLSLSSSKYRIBAA 1309
DB 2658 V---EAVKIKKTL-KNLTVTET-QDAVFTVELTHPNVKGQVQWIKNGVVLSEKVAISVK 2712

8764	EGQBY--KFRVIAKNKFCGPPVEIGP--ILAVDP-----	8799
6624	YGSEPFVSPBIIQONPVSEASDIWAMGVIYSLSITCSPFAGESDRATLLNVLGRVSW	6683
8795	LGPP--TSPE-----RLIYTERTKST-----ITLDMKEPRNGG	8826
6684	SPMAHLSE---DAKDFIKATIQAPQAPRPSAAQCLSHPWFLKSPAPAEAHFINTKQLK	6739
8827	SPIQGYIIEKRHDKPDF-----ERVNKRCLPTTSLFVENLDEHQMYFRVK---	8873
6740	PLLARSWORSLMSYKILVMRSIPELLRGPPDPSISJGVARHLCRDTGSGSSSSSSDNE	6799
8874	-----AVNEIGESEPSLP-LNVV-----TODDE	8895
6800	LAPFARAK-----SLP-----PSPVTHSPL-----LHPRGFLRPSASL---	6935
8896	VPPTIKLRLSVRGDTIKVKAGEPVHIIADVTGLPMPKLEWSKNETVIEKPTDALQITKEE	8955
6936	AEASERTEAPAPASPEGAGPPAAQCQVPRHSVIRSLFVHQAGESPEHGALLAPGRRHP	6995
8956	VSRSEAKTELISIPKAVREDKGTVTVTASNLGVSFRNVHVEVYDR-----PSPPRNL	9007
6996	ARRHLLKGGVIAGALP---GLREPLMEHRLVEEAREEQA-----TLLAK-----	6939
9008	AVTDIKAESCYLTDWAPLDNGGSE--ITHXVIIDKRDASRKKAWEVNTTAVEKRYGIWK	9065
6940	APSFETALRLPASGTHLAPCHSHLSHDSPSTPRPSSEACGEAQRLPSPAGCAPTRDM	6998
9066	LIPNGQYEFVRVANKY---GISDECKSDKVVITDP-----YRLPGPP--GKP-KVL	9111
6999	GHGQSGKOLPST-----GSHIPGTAQ--PERPSPDP--WGQAPAPCHPKQG-----	7044
9112	ARTKGLMSLVWTPPLDNGGSPITGYMLEKREBEGSPYMSVRAPITKVGLKGVEFNVPRL	9171
7045	EGCS-PPHVAVPCPPGSPFPGCKEAPLVPSSP-FLGQQAAPPAPAKASPPLDKXMGPG	7101
9172	LEGVKYOFRAWINAAIGPPSE-----PSPDEVAGDPIFFPGP--PSCPEYKDKT-KS	9222
7102	DISLPGRPKPGPCSSPGSASQASSQSVSLRVSSQVGTGPGPSLDAEGW--TOEAEDLS	7159
9223	SISLGWP-----PAKGG-----SPIKGYIVEMQEGT-----TDMKRVNEPDKLI	9264
7160	DS-----TPTLQRPQOATMRKSLGGRGGYAGVAGYGTGTFAGGDAGMLQOGPMWARI	7215
9265	TTCECVVNLKE-----LAKYRP-----RVK--	9285
7216	AVSQSEEBEQAERAEQSEEQEQAERAEPLQVARSAPVPEVG-----RAPT	7262
9286	AVNEAGESPSTDTGEIPATDIOBE-----PEVFI-----DIGAQDCLVCKAGSQRIPA	9335
7263	RSSPEPTP---WEDIGOVSLVQ---IRDLSGDAEAADTISLDISEVDPVYVNLSDLYDI-	7315
9336	VIGRPTPKSWEPPDGKAKAMKDGVDHIDIEDAQLETAENSSVIIPECKRSHTKGYSIT	9395
7316	KYLPPEFMFIKRVKPSAQPEPPSPMAEELAEFP-EPTWPWPEGLG-----	7360
9396	AKNKAGQKTANCRVKVMVDGPPKDKLVSDITRGSCRLSKMPPDDGGDRIKGVVIEKRT	9455
7361	-----PHAGL-----EITESEEDVALLAEAAVGRKKRWSSPSLSLFFHPCGRHL	7404
9456	IDGKAWTKVMPDCGTTTFVVPDLLSEQQYFFRVRAENRFGIGPPVETIQRTTARDP--IY	9513
7405	PLDEPAELGLRERVKASV-----EHLISRLTKGRPEGLEKEG-----	7440
9514	PPDPPIKLGITKNTVHLISWPKPKNDGGSPTHYIIVECLAWDPTGTKKEAWRQCNRKD	9573
7441	-----PPRKPKPLGSLFRSLGSLKSW-----DRAPTF-LRELSDET	7474
9574	VEELOFTVEDLVEGGEYEFVRKAVNAAGVKPSATVGPCDQCRQDMPDPPSIDLKEFME--V	9631
7475	VLGOSVTLACQVSAQPAQAATWSKDGAPLESSSRVLISATLKNQFLL--TILVVVAE---	7529
9632	BEGTNVNIKAKIGVPPPTLTWFKAPPKPKDNKEPVLTYDTHVNKLVVDDTCTVLIQOSRR	9691

Qy	7530	-DLGVYTCVSNALGTVTTTGVUKAERSSSPCPDIG-----EVTADGVLLVWKPVESY 7588
Dy	9692	SDTGLYITITAVNNLGTASKEMRLNVLGRPG----PPVGPIKFFSVSADQMTLSWFPKOD 9747
Qy	7584	GP---VTYIVQ-CSLEGGSWTTLASIDFDCYLTSLKLSRGTTYTERTACVSKAGMG-PYS 7638
Dy	9748	GGSKITNYVIEKREARKTWVHSSEPKECTYTPKLLGHEYVFRIMAQNKYGIGEPD 9807
Qy	7639	SPSE 7642
Dy	9808	SEPE 9811
RESULT 9		
Q10466 PRELIMINARY; PRT; 26926 AA.		
ID	Q10466	
AC	Q10466;	
DT	01-NOV-1996	(T-EMBLrel. 01, Created)
DT	01-NOV-1996	(T-EMBLrel. 01, Last sequence update)
DT	01-MAR-2004	(T-EMBLrel. 26, Last annotation update)
DE	Titin, heart isoform N2-B (BC 2.7.1.-) (Connectin).	
GN	Name=titin;	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Heart;	
RX	MEDLINE=96026330; PubMed=7569978;	
RA	Labelt S., Kolmer B.;	
RT	"Titins: giant proteins in charge of muscle ultrastructure and elasticity.";	
RT	Science 270:293-296(1995).	
RN	[2]	
RP	SEQUENCE OF 2277-25376 FROM N.A.	
RX	MEDLINE=92258380; PubMed=1582406;	
RA	Labelt S., Gautel M., Lakey A., Trinick J.;	
RT	"Towards a molecular understanding of titin.";	
RN	EMBO J. 11:1711-1716(1992).	
RN	[3]	
RP	SEQUENCE OF 1976-2014 FROM N.A.	
RA	Labelt S.;	
RL	Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.	
RN	[4]	
CC	CHARACTERIZATION.	
RX	MEDLINE=95331314; PubMed=7607248;	
RA	Gautel M., Castiglione-Morelli M.A., Pfuhl M., Motta A., Pastore A.;	
RT	"A calmodulin-binding sequence in the C-terminus of human cardiac titin kinase.";	
RL	Eur. J. Biochem. 230:752-759(1995).	
CC	-1- FUNCTION: THIS GIANT MUSCLE PROTEIN MAY BE INVOLVED IN MUSCLE ASSEMBLY AND IN MAINTAINING THE STRUCTURAL INTEGRITY OF SARCOMERES. MAY HAVE PROTEIN KINASE ACTIVITY.	
CC	-1- TISSUE SPECIFICITY: MUSCLE-SPECIFIC.	
CC	-1- SIMILARITY: TO THE CATALYTIC DOMAINS OF OTHER SERINE/THREONINE KINASES.	
CC	-1- SIMILARITY: BELONGS TO IMMUNOGLOBULIN SUPERFAMILY. CONTAINS 112 IMMUNOGLOBULIN C2-LIKE DOMAINS AND 132 FIBRONECTIN TYPE III-LIKE DOMAINS.	
CC	EMBL; X64698; CAA45939.1; -;	
DR	EMBL; X90568; CAA62188.1; -;	
DR	PIR; I38344; I38344.	
DR	PDB; 1BPV; NMR; @=14850-14961.	
DR	PDB; 1G1C; X-ray; A/B=2027-2125.	
DR	PDB; 1NCT; NMR; @=26050-26155.	
DR	PDB; 1NCU; NMR; @=26050-26155.	
DR	PDB; 1TIT; NMR; @=5244-5341.	
DR	PDB; 1TIU; NMR; @=5244-5341.	
DR	PDB; 1TKI; X-ray; A/B=24748-25068.	
DR	GO; GO:0005856; C:cytoskeleton; IEA.	
DR	GO; GO:0016020; C:membrane; IEA.	

DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0005516; F:calmodulin binding; IEA.
 DR GO: GO:0004896; F:hematopoietin/interferon-class (D200-domain...); IEA.
 DR GO: GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO: GO:0005198; F:structural molecule activity; IEA.
 DR GO: GO:0007517; F:muscle development; IEA.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR000382; Cytok receptor 2.
 DR InterPro: IPR000577; FGGY Kin.
 DR InterPro: IPR003961; FN III.
 DR InterPro: IPR008957; FN III-like.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003598; IG c2.
 DR InterPro: IPR003006; IG c2.
 DR InterPro: IPR011009; Kinase like.
 DR InterPro: IPR002016; Peroxidase.
 DR InterPro: IPR000719; Prot.kinase.
 DR InterPro: IPR002290; Ser thr_pkinase.
 DR InterPro: IPR008266; Tyr_pkinase_AS.
 DR Pfam: PF00041; fn3; 132.
 DR Pfam: PF00047; ig; 3.
 DR Pfam: PF00069; Pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00060; FN3; 132.
 DR SMART: SM00408; IGC2; 23.
 DR SMART: SM00220; S_TKc; 1.
 DR PROSITE: PS00933; FGGY KINASES_1; UNKNOWN_1.
 DR PROSITE: PS50853; FN3; 132.
 DR PROSITE: PS50835; IG-LIKE; 87.
 DR PROSITE: PS00290; IG MHC; UNKNOWN_1.
 DR PROSITE: PS00435; PEROXIDASE_1; UNKNOWN_1.
 DR PROSITE: PS00011; PROTEIN KINASE DOM; 1.
 DR PROSITE: PS00109; PROTEIN KINASE_TYR; UNKNOWN_1.
 KW Calmodulin-binding; Cytoskeleton; Immunoglobulin domain;
 KW Muscle protein; Phosphorylation; Repeat;
 KW Serine/threonine-protein kinase; Structural protein.
 FT DOMAIN 1370 1389 4 X 5 AA TANDEM REPEATS OF R-M-S-P-A.
 FT DOMAIN 4429 4614 GLU/LYS/PRO/VAL-RICH.
 FT DOMAIN 24731 25070 CATALYTIC.
 FT DOMAIN 25030 25056 CALMODULIN-BINDING.
 FT CONFLICT 22277 22277 T -> P (IN REF. 2).
 FT CONFLICT 22449 22449 E -> G (IN REF. 2).
 FT CONFLICT 22454 22454 T -> Q (IN REF. 2).
 FT CONFLICT 23324 23324 S -> L (IN REF. 2).
 SQ SEQUENCE 26926 AA; 2993428 MW; D5EEDC3254DF523 CRC64;

Query Match 8.7%; Score 3577.5; DB 2; Length 26926;
 Best Local Similarity 20.2%; Pred. No. 1e-105;
 Matches 1929; Conservative 1288; Mismatches 3662; Indels 2663; Gaps 320;

Qy 4 PQSGAPFLTRPKAFVVSVGKDATLSQIVGNPTQVSWEKDQOPVTAGARFRLA---Q 60
 Db 1030 PGEPAAPFYITKPVVKLVGGSVVFGCGVGNPKPHVYWKSGVPLTGTGRYKVSYNQ 1089

Qy 61 DGDYRLTLIDLALGDSQYVCRARNAIGFAFAVGL-----QVD 100
 Db 1090 TGEC- KLIVISMTFADDAGETYIVVRNKHGETSASALLSEADYBLLMKSQOEMLYQTQVT 1148

Qy 101 A---EAACAEQAPHL----- 113
 Db 1149 AFVQEPVEGTAPGFVSYEYKEYEKEQALLRKNAKDVVVVRYTVBDEPHISSFEERL 1208

Qy 114 -----LRPTSIRVREGSEATFCRV 133
 Db 1209 IKEIYRIIKTTBELLEEDGEERKMAVDISESEAVESGFDLRIKNYRILEGMGVTFFCKM 1268

Qy 134 GGSPPRAVSKDGRRLGEPDPVVRVEELGEASA-LRIARPRDGGTYEVRAENPLGA 192
 Db 1269 SGYLPKIAWYKDGKRI--KHGERYQMDFLQDGRASLRIPVVLPEDEGIYTAFAASNIKN 1326

Qy 193 ASAAALVDSDAADTASR--PGSTALLAHLQRR---REAMEAGCAPASPP----- 240
 Db 1327 AICSGKLYVEPAAPLGAFTYIPTLEPVSRLSFRSVSRPMSPARMSPARMSPARM 1386

Qy 241 -----STGTR-----TCTVTEGKHARLSCVYVTGEPKPTVWKK 273
 Db 1387 SPARMSERRLEETDESQLERLYKPVFLKPVSKLEGANCRFDLKVGVGPMPTETFWH 1446

Qy 274 DQQLVTGERRH--VVYEDAOENFVLKILFCQSRGLYTCTTASNLAVGQTVSSVLVWV--- 328
 Db 1447 DQQQIVNDYTHKVIKEDGTQSLI--IVPATPSDSGEWTVVAQNAGRSSISVILTVAV 1504

Qy 329 ----- 328
 Db 1505 EHQVKPMFVEKLVKNVNIKEGRLEMKVRATGNPNPNDIVWLKNSDIIIVPHKYPKIRIBGTK 1564

Qy 329 -----R 329
 Db 1565 GEAAALKIDSTVSQDSAWYATATANKAGDRTTRCKNVVEFAPEPBERKLIIPRGTYRAK 1624

Qy 330 EPAVP-----FKRLQDLVREKESATFLCE---VP 357
 Db 1625 EIAAPELEPLHLRYGQEWBEGDLYDKEKQKQKPFKKKLTSLRLKRFPAHFECLTPIS 1684

Qy 358 QPSTEAAWFKBETRL-----WASAKYGI-----EEGETERRLT 390
 Db 1685 DPTMVVEVLDHGKPLEAANRLRMINEFGYCSLDYVAYSRSDSGIITCRATNKYGTDTSA 1744

Qy 391 VRNVSADDDAVYICETPEGR-----TVAELAVQGNL-----LRKLPRKT 430
 Db 1745 TLIVKDEKSLVESQLPEGRKGLQRIEELERMAHEGALTGVTTDOKEKQKQPDIVLYPEPV 1804

Qy 431 AVRVDGTAMF-CVELAVPVGPHWLNRQEEVAVAGRVASAEGRTHLTISQCCLEVDVGQ 489
 Db 1805 RVLEGETARFCRVGTGYPQPKVNWYLNQLIRKSKRFRVYDGI-HYLDIVDCKSYDTGE 1863

Qy 490 VAFWAG-----DCQSTRFCVSAPRPDLPQ-----PVD 518
 Db 1864 VKYTAENPEGVIEHKVKLEIQORDFRSVLRRAPPEPRPEPHVHPGKLOFEVQKVDPRVD 1923

Qy 519 -----PVVKARMESSVILSWSPPHGE-----RPVITDGY-----LVKKKLGTYTW 560
 Db 1924 TTETKEVVKLKRAERTHEKVPESSELSKRFKRTTEGYEYAITAVELSKRKQDSYEE 1983

Qy 561 I-----RCHEABWVATPELTVD---VAEGNFOFRVSALNSFGQSPYLEPGTVHLPAPKL 613
 Db 1984 LLRKTQDELLHW--TKELTEEEKKALAEKGKITPTFKPKDIELSPSME-----APKI 2034

Qy 614 AVRTPLKAVQAVEGGEVTFESVDLTVASAG-----EWFLDQALKASS-VYEHCHDTRH 666
 Db 2035 FERIQSOTVG--QGSDAHFRVR-----VGKPDPECEWYKNGVKIERSDRIYVWYWPEDNVC 2088

Qy 667 TLTRIREVPASLHGAQLKF-----VANGIESIRMEVRAAPGLTANKPPAAAAAREVLARLHE 722
 Db 2089 ELVIRDVTA-BDSASIMVKAINIAGETSSHAFLLVQAKOLITFTQ-----ELQDVVAKED 2143

Qy 723 E-AQLLAELSDQAAAVTWLKGRTLSFGPKYFVQASAGRVLVLRDVARDDAGLYECV-- 779
 Db 2144 TWATFECSETSEPPVKVYKQDGMVHEGDKYRMHSDRKVHFLSILTIDTSDAEDYSCVLV 2203

Qy 780 -SRGRIAYQLSVOGLARFLHKDMAGSCVDVAVAGGAQFECETSEAHVHVHYKDGMELG 838
 Db 2204 EDENVKTTAKLIVEGAVVFRVVKELQD--IEVPESYSGELECIIVSPENIEGKWHNDVELK 2261

Qy 839 HSGERFLQEDVGRHRLVAATVTRQDEGTYSVCRVGEDSVDFRLVRVSEPKVVFVAKQLARR 898
 Db 2262 SNCKYITISRRG-RQNLTVKDVTKEDQGEVSVFIDGKTKTKLKM-KPRPIALQLGSLDQ 2319

Qy 899 KLOAEAGASATLSCEVAQAQTEVTYKDGKLSKSSKVCWMEATGCTRRLLVQQAQADAG 958
 Db 2320 KVC--EGDIVQLEVKVLSLESVEGVWMDQGEVQPSDRVHIVIDKQSHMLLIETDKEDAG 2377

Qy 959 EYS-----CEAGGQGLRSLFDLVKPKVVFVAKQDVAVHSEVQAEAGANATLSCEVAQAQ-- 1010
 Db 2378 NYSFTTIPALGLSTGRSVSYSDVITPL-----KDVNVIETKAVLECKSVDPDV 2427

Qy 1011 AEVMMYKDGKLSLKVHVEAKGRRRLVVOQAGKTADAGYSCBQGVSPRLHITEP 1070
Db 2428 TSVKWLNDGQIKPPDRVQAIKVGTKORLVNRTSHASDEGPKYKLVGVETNCLNSVEKI 2487
Qy 1071 KQMPAKEQSVNEVQAEAGASAMLSCEVAQAOTEVTVYKDGKLSLSSSKVMEKVGCTRR 1130
Db 2488 KIIRGLDLTCTEQ-----NVVEVELSHGIDVLNFKDKKEIKPSSKYIEAHGKIYK 2542
Qy 1131 LVLPOAGKADAGEYSCBQGVSPRLHITEPQVFAKEQSVHNEVQAEAGTTAMLSCV 1190
Db 2543 LTVLNMKDDGKTYFVAGENMTSGK--TVAGGAISK--PLTDQTVAES--QEAUFCEV 2597
Qy 1191 AOPOTEVTVYKDGKLSLSSKVRMEVCKTRRLVVOQVQKADAGEYSCBQGVSPRLHITEP 1250
Db 2598 ANPDSKGEWLDGKHLPLTNIRSESDGHRLLIAATKLDIIEYTYKATSKTSAKL 2657
Qy 1251 ITEPKAVFAKEQLVHNEVTRTEAGASATLSCEVAQAOTE-VTVYKDGKLSLSSSKVRIBAA 1309
Db 2658 V---EAVKIKKTL-KNLTVTET-QDAVFTVELTHPNVKGQVWIKNGVLESNEKVAISVK 2712
Qy 1310 GCMQLVVOQAGADAGEYTCBAG--GORLSFHLID-----VSEPKAVFAKEQLAHRKQAE 1363
Db 2713 GTIYSLRIKCAIYVDESUYGFRGLGASARLHVETVKIIEKPKDVTALEN----- 2763
Qy 1364 AGATATLSCEVAQAQ--TEVTVYKDGKLSLSSSKVRMEAVCTRRVVOQACQADTCYS 1421
Db 2764 ----ATVAFESVSHDTPVFMFKHSVEIKSPDKHRLVSRKVKHMLQNTSPSDAGEYT 2819
Qy 1422 CEAGQRLSFLSDVAEPKVPFAKQPVH-----REVQAQAGASTTSCCEVAQAOTEVVMWY 1476
Db 2820 AVVQLECKAKL.FV-----ETHITKTKMNIIEVETPTASFECEVSHFNVPSMWL 2869
Qy 1477 KDGKLSLSSSKVRMEAVGCTRRVVOQAGADAGEYSCBQGVSPRLHITEPQVFAKEQSVHNEV 1536
Db 2870 KNGVEIEMSEKFKIYVQKHLQIIMNTSTSDSAEYTFVCGNDQVSAATLTVT-PIMITS- 2927
Qy 1537 EQPASREVQAEAGTATLSCEVAQAOTEVTVYKDGKLSLSSSKVRMEAVGCTRRVVOQAE 1596
Db 2928 ---MLKQINAEKOTITFEVTVNYEGISYKWLKNGVEIKSTDKQMRKTLKTHSLNIRV 2984
Qy 1597 QOADAGEYSCBQGVSPRLHITEPQVFAKEQSVHNEVQAEAGASATLSCEVAQAOTE 1656
Db 2985 HFGDAADYTFVAGKATSTATLYVEARHIEFRKH---IKDIKVEKKRAMFECEVSEPDIT 3041
Qy 1657 VTVYKDGKLSLSSSKVRMEAVGCTRRVVOQAGADAGEYSCBQGVSPRLHITEPQVFAKEQSVHNEV 1716
Db 3042 VQMKDDQELQITDRIKIQEKYVHRLIIPSTRMSDAGKYTWAGGNVSTAKLFVEGRDV 3101
Qy 1717 QISERPCCREPLVKEHEDIIILTATLTPSAATVTLKDGVEI--RRSKRHETASQGDTH 1774
Db 3102 RI--RSIKKEVQVIEKQRAVV--EFVNEDDVDAAHYKDGIEINFPQQRHKYVYVERRIH 3157
Qy 1775 TLTVHGQVLDLSAIVSCRVGAEGQDFPVQVB-EVAAKFCRLLEPVCGLGTVTLACELS 1833
Db 3158 RMPITSETROSDAGEYTFVAGNRSSVTVYNAPPEPPQVLOLQPTVVOG-----K 3208
Qy 1834 PA--CAEVVWRCGNTQPRV-----GKRFQVMAEGRVSLTVLGLRAEDAGEYVC 1890
Db 3209 PARFCAMI--SGRPQPKISYKBEELQSTGFKCKFLHDGQEYTLILLIEAFPDAVYTC 3265
Qy 1891 ESRDDH-----TSAQITVSVPRVK-----FMSGLTVNAEAGEATFQCVSPS 1925
Db 3266 EAKNDYGVATTASLSVEVPVSDQEMPVYPPAIITPLQDVTSEQPARFQCRVSGT 3325
Qy 1926 DVAVVWFRDQALLQFSEKFAISQSGASHLTIISDLVLEDAGQITVEABGA-----SSSAAL 1981
Db 3326 DLKVSWSKDKKIKPSRFRMTQPEDTYQLIEAAYPEDEGTYTFVANNVGVQSSTANL 3385
Qy 1982 RVREAPVFLKKLBPQ-----TVRESSSVTLEVELTR----- 2013
Db 3386 SLEAPESILHERIQEITEMEMKEFSSFLSABEELGSAELQLSKINETLELLSESPPV 3445
Qy 2014 -----PWPELWRNATAPACKNVEI 2035

Db 3446 TKPSEKEGTGPIPIKEVSNADISMGDVATLSVTVIGIPKPKIQMFFNGVLLTPEADYKP 3505
Qy 2036 HAEGARHRLVJHNVGVFADRGFFGCTPDD-----KT 2066
Db 3506 VFDGDDHSLIILFTKLEDEGEYTCWASNDYKTKICSAYLKINSKGEKHOTETESAVAKS 3565
Qy 2067 QAKL-----TVMEQVRLVRGLQAV-EARBOGTATMEVQLSHADVDGWSRTRGLRFQ 2117
Db 3566 LEKLGPCPPHFLKELKPIRCAOGLPAIFETVVVGEPAPTV-----TWFKENKQL- 3615
Qy 2118 QGPTC---HLAVRGP--MHTLTLSGLRPEOSGLMVFAKEGV---HTSARLVW----- 2161
Db 3616 ----CTSVYVYIIHNPNNGSGTFIWNDFQREDSGLYICAEENMLGESTCAEELLVLEBTD 3671
Qy 2162 -----TELPVSFSR-PL-----QDVVVTTEKBEKVTLECELSRPNVDVRW 2198
Db 3672 MTDTPCKAKSTPEAPDFPOTPLGPAVEALDSQEIATFVKDTILKAALITEENQQLSY 3731
Qy 2199 -----LKDGVELRAGTMAIAAQ-----ACRSLTY----- 2225
Db 3732 EHIKANELSSQLPLGAQELQSILOQDKLTPESTREFLCINGSIHFOPLKEPSNLOLQI 3791
Qy 2226 ---RCEPADQGVVYCDADAQOS----- 2244
Db 3792 VQSKTFSEKILMPPEPETQAVLSDTEKIPPSAMSIEQINSLTVEPLKTLAEBEGNYP 3851
Qy 2245 SASVKVOGRTY-----DAGEIQVFA---ENAESRAQLRVKE----- 2287
Db 3852 QSSIEPPMHSYLTSVAEBVLSEKVTYSDTNRQORVTLQKQEAQSALILSQSLAEGHVES 3911
Qy 2286 -----LPTVLRPLRDKIAMEKHGV----- 2308
Db 3912 LQSPDVMIQVNYBPLVPSEHSCTEGGKILITESANPLENAGQDSAVRIEESKSLRFLPAL 3971
Qy 2308 ----- 2322
Db 3972 BEKQVLLKEHSDNVNMPDQIIESKREPAVKVQVQGRDLILSKESLLSGIPBEQRLN 4031
Qy 2309 LECQVSR-BAQV-----RW----- 2322
Db 4032 LKIQICRALQAASVASEQPLGFSEWLRNIEKVEAVNITQEPRIHMCMLVLTSAKSVT 4091
Qy 2323 -----FKGSQ 2327
Db 4092 VTIIEIDVDPQMANLKMELDALCAIIEEDILTAEGRIOQAKTSLOBEMDSFSGS 4151
Qy 2328 ELQPGKVELVSDGLYRKLIIISDVHAEDEDTYTCADGV----- 2366
Db 4152 KVEPITPEVESKYLIIISTEVSFYVQSRVKY-LDAPTVTKGVASAVVSDKQESLXPS 4210
Qy 2367 ----KTSAQPFVPEQOSITIVRGLQDVTVMPEPAPA-----WPECETSIPIPSVR 2408
Db 4211 EEKESSESSESTEEVATVKIQEAGGLIKEDGPMIHTPLDVTVSEGDIVHLTTSITNAK 4270
Qy 2409 PPKWLLGKTVLQAGNVGLBQEGTVHRLMLRR-----TCSTMTGPHVFTVGKRS 2458
Db 4271 EVNMYFENKLVPSDEKFKCLQDQNTYTLVIDKVNTEHDQGEYVCEALNDS-----GKTAT 4325
Qy 2459 SARL-VVSDIPVLTPLPEKPTGRELQSVLSCDFRAPKA-VQWYKDDTPLSPSEKPKM 2516
Db 4326 SAKLTVVKRAAPVKIKIEPLEVALGHAKFTCEIQSAPNVRQWFKAGRIYESDKCSI 4385
Qy 2517 SLEQMAELRILRLMPADAGVYRCQAGSAHSSTEVTVEAREVTVTG-----PLQDA 2567
Db 4386 RSSKYISLELRLQVVDCCGEYTCASNEYGSVCTA-TLTVTPGGEKKYRKLPERKP 4444
Qy 2568 EATEEGWASFCESLSEDEEVEWLSNGMPLVNDSFHIEHSHGRHTLVLSIQRADAGIV 2627
Db 4445 EPKEE--VWLKSVLRKRPEEPEKPKCL-----EKVKDPAVPEPPPKPVEEVETV 4497
Qy 2628 RASLSKYSTSARLEVRVKPVVFLKALDLSABEERGLTALQCEVSDPE-----AHVV 2678

Db	4498	TKRERKIPEPTKVP-EIKPAIPLPAPEPKPKPEAEVKTIKPPPVPEPTPIAAPVTPVW	4556
Qy	2679	WRKQVOLGSPDKYDFLHTAG-TRGLVVDVSPDAGLYTCHVGS-----BETRARVRV	2731
Db	4557	GKKAERK--APKEBAAPKPGIKVPPKKTSPIEAERRKLRPGSGGKPPDEAPFTYQL	4613
Qy	2732	HDLVGIGITRKLTMVBLE-----GESCSFECVLSHESASDPAMWTVGGKTVSGSSRFRQATR	2787
Db	4614	KAVPLKFVKEIKDIIILTESEFVGSSAIFACLVSPTAI--TTWKDGSNIRESPKHFIA	4671
Qy	2788	QGRKILVVRREAPSDAGEVFSVR-----GLTSKASLIVRBRPAAILKPLSDQ-WVADGE	2842
Db	4672	DKOKRKLHIIDVQLSDAGEYTCVLRLGNKEKTSTAKLVBEELPVRVFKLTBEEVTVVKQG	4731
Qy	2843	DVELRCELSRAGTPTVHMLKDRK-AIRKSOKYDVWCEGTMAMLVIRGASLKDAGEYTCBEV	2901
Db	4732	PLYLSCELNKE-RDVWRKDKIVVEKPGRIVPGVIGLMRALTINDADDTAGTYTVTVE	4790
Qy	2902	ASK-----STASLHVEEKANCFTTELTNIQVBEKGTAFTCKTEHPAATVTVWRKGLLELRA	2957
Db	4791	NANNLECSCKVVEVIRDLVKPIRDQHVKPKGTAFACDIAKOTPNIKWFKGYDEIPA	4850
Qy	2958	--SGKHQPSQGLTLRLTISALEKADSDTYTCDICQAOSSRAQLLVQGRVHIIEDLEVD	3015
Db	4851	EPNDKTEILRDGNHLYLKIKNAMPEDIAEYAVEIEGKRPYAKLTIGEREVELLXPIEDVT	4910
Qy	3016	VOEGSSATFRCKRISPANYEPVHWFLDKTPHLANELNEIDAOPGGVHVLTQLALKDSGT	3075
Db	4911	IYKESASFDAEISEADI-PCQWKLKGBLLPSPTECEIKAB-GGKRFITLHKVKLDQAGE	4968
Qy	3076	IYFEAGDQASAAALRVTEKPSVFSBELTDTATITBEGEDTLVCETSTCDIPMCWTKDQKTL	3135
Db	4969	VLYQALNAITTAITLVKIEIELDFAVPLKDVTPERRQARFEC-VLTREANVIMSGPDII	5027
Qy	3136	RGSAQCQLSHGHRAQLLITCATLQDSRYKCEAGCAGSSSIVRVHARPVPFQALKDLE	3195
Db	5028	KSSDKFDIIADQKGHILVINDSQFDEGVYTAEVGKKT SARLFTVGIKLFKMFPLEDQT	5087
Qy	3196	VLEGGAAATLRCVLSVAAPVKWCYGNVNLRPQDKYSLRQEGAMLELVNRLRPQDSGRYS	3255
Db	5088	VKEGETATFVCELSHEKMHVVWFKNDAKLHTSRVTLISSEKTHKLEMKVETLDDISQIK	5147
Qy	3256	CSFGDQTTSATLTVTALPAQFIGKLNRNKEATGEGATATLRCBSLKTAPVWRKGSSETLRDG	3315
Db	5148	AQVKELSTAQLKVEADPYFTVKLHDKTAVEKDBITLKCEVSKDVPVKWFKDGEIIVPS	5207
Qy	3316	DRYCLRQDAGMCELIQIRGLAMVDAEAYSCVCGERTSASLIRPMPAHFIRLHRQHSIE	3375
Db	5208	PYISIKADGLRRIILIKKADLKDKGEYVCDGCTDKTKANVTVEARLIEVEKPLYGVEFV	5267
Qy	3376	GATATLRCELS-----KAAPVWRKGRSIRLDGRHSRLQDQGAVCCELOICGLAVA	3425
Db	5268	CETAHFEIELESDPDVHGQWKLKGOPLTASPCCEIIEDGKHI-----ILLHNCQLGM-	5319
Qy	3426	DAGBYSCVCGERTSATLTVKALPAKFTTEGLRNEEAVEGATAMLWCELSKVAPV-EWRKG	3484
Db	5320	-TGEVVSFOAANAKSAANLKVRELPLIFITPLSDVKVFEKDEAKPECEVSRBPKTFRWLKG	5378
Qy	3485	PENLRDGRYTLROEGTFCCELOICGLAMADAGEVLCVCGQERTSATLIRALPARFTEDV	3544
Db	5379	TOEITGDDRFELIKDGTKHSWIKSAAPFEDBAKYNFEABDHTSGKLIIEGIRLKLFTPL	5438
Qy	3545	KNQBARCAGATVLOCEL-NSAAPVWRKGSSTLRDGRYSILRQDGTCKELQIRGLAMADT	3603
Db	5439	KDVTAKESAVFTVELSHDNIRVKVFNQDQRLHTRSVSMQDEBKTHSITFKDLSIDDT	5498
Qy	3604	GEYSCVCGERTSAMLTVRALPIKFTTEGLRNEEATGEGATVLRCELSM-APVBNWKGHE	3662
Db	5499	SQIRVEAMGMSSEAKLTVLEGDPYFTGKLDYTVGVEKDEVILQCEISKADAPVKWFKDGK	5558
Qy	3663	TLRDGRHSIRODGCARCELIQIRGLVAEDAGVLCWCGKERTSAMLTVRAMPSKEIEGLRN	3722
Db	5559	EIKFSKNAVIKTDGKRMLIUKKALKKDIGYTCCDGTDKTSGKGLDIEDREIKLVRPLHS	5618
Qy	3723	EE-----	3724
Db	5619	VEVMEETARFETISEDDIHANWKLKGEALLQTPDCEIKEEGKIHSLVHNCRLDQTGG	5678
Qy	3725	-----ATBGDTATLWCELS-KAAPVWRKRGHETL	3752
Db	5679	VDFQAAANVKSASHLVRKPRVIGLLRPLKDVTVTAGETATFDCELSYEDIPVEMWLKGGKL	5738
Qy	3753	RDGRHSLRODGRSCELIQIRGLAVVDAGEYSCVCGQERTSATLTVRALPARFIEDVKNOE	3812
Db	5739	EPSDKVVRSEGKVHTLUTLRDVKLEDAAGEVOLTAQDFKTHANLFPVKPPVEFTKPLEDQT	5798
Qy	3813	AREGATAVLOCELSKA-APVWRKGSSTLRDGRYSILRQDGTCKELQIRGLSVADTGBYS	3871
Db	5799	VEEGATVLECEVSRENAKVKFKNGTEILKSKYEIVADGRVRLVHIDCTPEDIKTYT	5858
Qy	3872	CVCQERTSATLTVRAPQVPRPQLSQLOABEGSTATLQCELSSEPTTVVWVSKGLOLQA	3931
Db	5859	CDAKDFKTCNLTNVPVPHVBFRLPLTDLQVREKEMARFECELSRENKVKWFKDGBAIKK	5918
Qy	3932	NGRREPLQCTAELVLQDLQREDTGYTCTCGSQATSATLTVTAAPVRLRELQHOEVD	3991
Db	5919	GKTYDIISKGAVRILVINKCLLDDAEYSCVTRTARTSGMLTVLEEEAVFTKNLANLEVS	5978
Qy	3992	EGGTAHLCCELSRAGASVEWRKGSLLQFPKCAKYQMVQDGAALLLVRGVEQEDAGDYTC	4051
Db	5979	ETDTIKLVCEVSKPGAIEVWYKGDDEEIIETGRVEILLTEGRKILVIONAHLEDAAGYNCR	6038
Qy	4052	TGHTQSMASISVRVPRPKFTRIQLSLEQETGDIARLCCQLSDABSAGVWQWKEGVELHA	4111
Db	6039	LPSSRTDGKVKVHELAAAEFISKPNLEILSEKAFAEFVCSIS--KESPPVQWKRDKDTLES	6096
Qy	4112	GPKYEMBSQCATRELLIHOLEAKDGTGEYACVTCGQKTAASIRVTEPEVTVIRGLVDAEVT	4171
Db	6097	GDKYDVJADGKRVLVVKDQATLQDMGYTVVWGAARAAHLTVIE-KLRIVVPLKOTRVK	6155
Qy	4172	ADEDEPFCVSRAGATGVQWCLQGLPLQSNVEVTEVAVRDRGRIHTLRKGVTPBEDAGTVS	4231
Db	6156	EQGEVFNCEVNTGAKAKWFRNEEAIFDSSKY--IILQKDLVYTLAIRAHLDDQANYN	6213
Qy	4232	FHLGNH-----ASSAQITVRAPEVTTILEPIQDVOLSEGQASFOCLRSLRSGQBARWALG	4286
Db	6214	VSLTNHNGENVKSAANLIVEEEDLRIVEPLKDIETMEKKSVTFWCKNRLN-VTLKWTQN	6272
Qy	4287	GVPLQANEMNDITVEQGTLLHLTLHKVTLLEDAGTVSPHVGTCSEAOIK-VTAKNTVVVG	4345
Db	6273	GEVPPPNRVSVYRDK-YKHLMTIKDCGFPDEGEYIVTAGQDKSVABELLIEAPTEFEVH	6331
Qy	4346	LENVEALEGCEALFECOLSQPEVAAHTWLLDDBPVRTSENAEVVFPFENGLRHLLKRLR	4405
Db	6332	LEDQTVTEFDVAFVSCQLSR-EKANVKWYRNGREIK--EGKYYKFEKDGSHRLIHKDCR	6388
Qy	4406	PQDSCRVTFIAGDMVTSAFITVGRWRLIELEPLKNAAVRAGAAQARFTCTISEAVPVGEAS	4465
Db	6389	LDDECEYACGVEDRKSRLRFVEIPIRPPQDILEAPGADVVFIAELNKK--DKVEVQ	6446
Qy	4466	WYINGAAVQPDSDWTVTADGSHQALLRSAPHPHAGEVTFACRDAVASAR-----	4516
Db	6447	WLRNMVVVQGDKH-QMMSGKTHRLQICDKPRDQGEYRFFIAKDKARAKLELAAAPKI	6505
Qy	4517	-----	4516
Db	6506	KTADQDLVVDVGKPLTMVVPYDAYPKAAEAWFKENEPLSTKTTIDTTAEQTSFRILEAKKG	6565
Qy	4517	-----LTVLGLPDPPEAEVVAHSHHTVTLSWAAPMSDGGGLC	4555
Db	6566	DGGRYKIVLQNKHGKABGFNLKVIDVPGVFRNLEVTETFDGEVSLAWBEPFLDTGSKII	6625
Qy	4556	GYRVEVKEGATGOWRLCHELVPGPECVWDGLAPGET-YRPRVAAGVPVGAPEVPHLPQTV	4614
Db	6626	GYVVERDIRKKTWVLATDRAESCEFTVTLQGGVEYLFVRSARBNRVGTGEPVETDNPV	6685

QY 4615 RLAPPPKVPPOP-----SAPESRQVAAGEDVLSLEVVAAEAGEVINHKGM 4660
Db 6686 E-ARSKYDVGPPLNVTITDNRFGVSLTWPEPYDGGAAETNYVIELRDKTSIRWDTAM 6744
QY 4661 ERIQGGREFEVVSGRQOQMLIKGFTABDOCEYHCGLAQSGICPAAATFOVALSPASVDE 4720
Db 6745 T-----VRAEDLSATVTVVSGOYSFRVRAQNRIGVCK-----PSAATPFVKVADPIERP 6795
QY 4721 APOPSLPEAAQOGLHLLBALARKRMSREPTLDSISELPEEDGRSQRLOPQAEAEVAP 4780
Db 6796 SPPVNLTSDDQSSQVQLKWEPLKD--CGSPILGYIIEERCEE-GKDNWIRCNM-KLVP 6850
QY 4781 DLSGYSSTADELARTGDADLSHTSDDESRAGTSLVLYLKAGRPGETSPILASKVGAPAA 4840
Db 6851 ELTYKVTGLEK-----GNKLYRVAENKAGVSDPSEI-----6883
QY 4841 PSVRPQOOQELAAVRPPLGLDSTKD-LGDPMSMDKAAVK-----IQAAFGYK 4887
Db 6884 -----LGPLTADDAFVEPTMDLSAFKDGLEIVPNPILVPSTGY- 6924
QY 4888 VRKEMKQOEGMFHSTFGDTBAQVGDALRLBECVASKADVRAWMLKOGVELTIDGRHHHID 4947
Db 6925 -----PRPTATWCFGDKLETGDRVKMKTLSA-----6951
QY 4948 QLGDGTCSLLIAGLDRADAGCYTCOVSN-----KRGQVTHSA 4984
Db 6952 -----YAELVISPSERSDKGIYTLKLENRVKTIISGEIDVNVIAAPSAPKELKFGDITKDS 7006
QY 4985 CVWVSGSEASESSGGELDDAFRAARRLRLFRK-----5021
Db 7007 VHLT-----WEPDDGGSPLYGYVVEKREVRKTKWMDFTDLEFTVPDLVQCKEYLF 7062
QY 5022 -----SPAESVDEELFSLADEGAPPEPADQTYREDBHFI-----5058
Db 7063 KVCARNKCGPEPAV-DEPVNMSTPATVPDPENVKWRDRTRANSIFLTWPPKNDGGS 7121
QY 5059 -----CIR-----FEALTEARQAVTRFQB--MFATLGIGVEIKLVEQGRPRVEM 5100
Db 7122 IKGVIVERCPRGSKWACGEPVATKMEVTGLEEGKYA-----YRVKT 7166
QY 5101 CISKETPAVVPPEPLPILLSDAAPVLTB--LQOQEVQDGYPVSCVVTGQPMPSVR 5158
Db 7167 LNRQASRSPRTEIEQAVDTQEAPEIFDVKLLAGLTVKAGTKIEPATVIGKEPKIT 7226
QY 5159 WFKQGLLEEDHYMINEDQOQHOLIITAVVPADMVGYRCLAENSMGVSSTKAEALRYDL 5218
Db 7227 WTKADMILKQDKRITI-ENVPKSTVTIVDSKRSDTGYIIEAVNVCGRATAVVEVNV-L 7284
QY 5219 TSTDYDTAADATESSSYFSAOGYLSRSREOGETSTDEGQLPQVVEELRLQVAPGTRLA 5278
Db 7285 DKPGPPAFTDVTNESCLLTWNPDRDDGS-----7316
QY 5279 KFLQVKGYAPAPRLWFKDQGLPTASAHIRMTGKILHTLEIISVTRSDSGOYAAVISA 5338
Db 7317 -----KITNYVER-----RAITSEVWHKL-----STVKTDTNFKATKLIPN 7353
QY 5339 MGAAYSSARLLVRGDEPEERKPSADVHQLVPPRMLERFTRPKVKXKS-SITFSYKVE-- 5395
Db 7354 KEYIFRVAENMYGAGEPVQASPIAKYQFOPPGPPTRLEPSDITKDAVTLTWCEPDDG 7413
QY 5396 GRPVTWHLREABRGVLWIGPDTPGYTVVASSAQOHSLLVDGRHQGTY--TCIASN 5453
Db 7414 GSPI-TGYVW-----BR-----LDPDQDKWVRNCNMPKVDITVRVKGLTNKKYRFRVLAEN 7464
QY 5454 AAGQALCSASLHVSLP-KVEEQEKVKALISTFLQG--TTOAISAOGLTASPADLGGQ 5510
Db 7465 LAG-----PKPSKSTEPILIKOPIDPPPPGKPTVKDVKTSVR-----LNWT 7508
QY 5511 KREEPLAAKEALGHLSLAEVGTETPLKLTQSITEMVSAKITQAKLQVPG-GDSDEDSKT 5569
Db 7509 KPEHDGAKIESYVYIEMLKGTGDENVR-----VAEGVPTTQHLPLGLMEGQEYSFR 7559
QY 5570 PSASPRHGRSPSSSIQSSSESESGDARGEIPDIYVVTADYLPILGAEQDAITLREGQY- 5628

Db 7560 VRAVNKAGESEPSB-----PSDPVLCREKLYP 7586
QY 5629 -----VEVLDAAH---PLRWLVRTKP-----TKSSPSRQGM--VSPAYLDRLK 5667
Db 7587 PSPRWLEVINITKNTADLKWTVPEKGGSPITWVIVEKRDVRRKGMQVTDVTKTKCT 7646
QY 5668 LSP--EWG-----AAAPPEPFGAEVSEDEYKARLSSVIOELLSEQAFVBEQLQSHH 5719
Db 7647 VTPLTEGSLYVFRVAEE-----NAIGQSDY-----TEIEDSVLAKOTFT-----7685
QY 5720 LQHLERCPHPIAVAGQKAVIFRNVVDIGRPHSHSFLOEQOCDTDDDDVAMCFI-KNOAAF 5778
Db 7686 -----TPGPPYALA-----VVDVKHVDLKWEPKPKNDGGRPIQRYVIEKERLG 7730
QY 5779 EOYLEFLVGRV-----QAESVVVSTAIQFYKKYAE-----BALLAGDPQ 5819
Db 7731 TRWVK--AGTAGPDCNFRVTDVIEGTEVQ--FQVRAENEAGVGHPSBPTILSIEDPTS 7786
QY 5820 PPPPLQHYLE-----QOVER-----VQRYOALLKELIRNKA 5851
Db 7787 PPSPLDLHVTDAGRKHIAIAWKPKPKNGGSPIIIGYHVEMCPVGTCKMVRVNSRPIKDLK 7846
QY 5852 RNRQNCALLEOAYAVVSALPORAENKLHVS-----LMENYPG-----TLEALGEPIRQG 5900
Db 7847 FKVEGVVPDKEY-----VLRVAVNAICVSESEISENVAVKDPCKCTID-----LETH 7897
QY 5901 HFIWEG-----APGARMPWKGNHRV-----FLFRN-----HLVICKP-RRDS 5938
Db 7898 DIIVIEGKLSIPVFPRAVPVPTVWHKQKGVKASDRLTMKNDHISAHLEVPKSVRADA 7957
QY 5939 RTDVTSVYFRNMKLSIDLNDQVEG-----DRAFEV-----WQ-----5973
Db 7958 GIYTIIT--LENKLGASATASINVKVLGPGCKDIKASDITKSSCKLTWEPPEFGGTPI 8015
QY 5974 -----EREDSVKYLLOARTAIKSSW-VKEIC-----GIQORLALPV- 6010
Db 8016 HVLREAGARTYIPVMSGENKLSWTVDIIPNGEYFPRVKAVNKVGGGEYIELKNPVI 8075
QY 6011 -----WRPPD--FEEELADCTAELGETVKLACRVISWYKDGKAVQVDPHHILIE 6064
Db 8076 AODPKQPPDPVDEVHNPTAE-----AMTITWKPPL--YDGGSKIM--GYIEBK 8121
QY 6065 DPDGS-----CALILDSLTCVDSG-----QWCFPAASAGNCSTLGKILVQVPP-RF 6110
Db 8122 IAKGEERWKRCNEHLPIITYTAKLEGEGEYQFRVRAENAAG-----ISEPSRATPPTKA 8177
QY 6111 VNKVRASPV-----EGEDAQFTCTIECAPYQIRWYKD-----CALLTT 6150
Db 8178 VDPIDAPKVILRTSLEVRGDEIADIASISGPIPTITWIKDENVIVPEEIKKRAAPLVR 8237
QY 6151 GNKFQ-----TLSEPRSGLLVLVIRAASKEBGLGYECBLVNLRSARASA 6195
Db 8238 RRKEVQEEEPFVPLTQRLSIDNSKKGESQLRVRDSLRPDHGLYMIKVENDHGIKAPC 8297
QY 6196 ELRTQS-----PMLQAQEOCHREQLVAVE-----DTTLERAOQE-----VT 6232
Db 8298 TVSLDTPGPPINFVEDIRKTSVCKWEPLDDGGSEIINYTLEKKDKTKPDSEWIVVT 8357
QY 6233 SVLK-----RLLGPK-----APGS-----TGLDTGPGPCP- 6258
Db 8358 STLRHCKYSVTKLIEGKEYLFRVAENRFGPGPCVKPLVAKDPFGPPDAPDKPIVEDV 8417
QY 6259 -----RGAPAL-----6264
Db 8418 TSNSMLVKMNEPKONGSPILGYWLEKREVNSTHMSVNSLLNALKANVDGLLEGITYVF 8477
QY 6265 ---QETGQPVTGTSEAPVPPRPVQPLLHEGGEQEPEAIARAQEWTVPIRMEGAAPG 6321
Db 8478 RVCAENAAGP---GKFSPPSPDKTAHDIPSPGP--PIPRVTDTSSTSTIELEWEPAPNG 8532
QY 6322 AG-----TGELLWD-----VSHVVRITQ-RITYTQAIDTHTA-----6354

Db 8533 GGEIVGVYFVDKLVGNTKNSRCTEKMIKVRQYTVVKEIREGADYKURVAVNAAGPPGE 8592
Qy 6355 -----RPPSMQVETIE-----DVQAQTGGTAQAEIIEGDPQSPSTWYKDSVOLVDST 6401
Db 8593 TQPVTVAEPOEPFAVELDVSKGGIIMAGKTLRIPAVVTGPPVTKWTKKEGSL-DKD 8651
Qy 6402 RLSQOQEGTYSVLVRHVASKADAGVYTCLAQNTGGQVLCARLLVLGGDNEP-----DSEK 6457
Db 8652 RYVIDNVGTGKSLIIKDARLDKHGRYVITATNSCGSKFAAARVEVF-----DVGPPVLDLKP 8708
Qy 6458 QSHREK--LHSPVEYKBEIGRVFGVRVQVHKGKILCAAKFPLRSRTAQAVERDI 6515
Db 8709 VVTRAKMLNDSDEGGSEITGFI--IERKAKM----- 8743
Qy 6516 LAALSHPLVTGLDQFETRKTLILILELCSSEELDLRYKGVVTEAEVKVYIQOLVREL 6575
Db 8744 -----HTWRQPIETERS-KCDITCLLSGQ 8766
Qy 6576 HYLHSHGV-----HLDIKPSNIMLVHPAREDIKICDPGFAQNTPAELOFSQVGS 6626
Db 8767 EY--KFRVIKAKFKCGPPVEIGP--ILAVDP-----LGP 8797
Qy 6627 BEFVSPEIILQNPVSEASDIWAMGVISYLSLTCSPPFAGESDRATL-LNVLEGRVSNSSP 6685
Db 8798 P--TSPE-----RLTYT-----ERQRTITLWKEPRSNNGSP 8828
Qy 6686 MAALHSE-----DAKDFIKATLQAPQAPSAQCLSHPWFLKSPAEBAHFINTKQKFL 6741
Db 8829 IQGYIIEKRHDKPDF-----ERNKRLCPTTSLVENLDEHQMYEVRK----- 8873
Qy 6742 LARSWQSLMSYKSLIWNRSIPELLRGPDPSPSLGVARHLCRDTGGSSSSSSSSNDELA 6801
Db 8874 -----AVNEIGSESESLP-LNVV-----IQDDEVP 8897
Qy 6802 PPARAK-----SLP--PSPVTHSPL-----LHPRGLRPSASL-----PREAE 6837
Db 8898 PTIKLRLSVRGDTIKVAGEVHPADVTGLPMPKIEWSKNETVIERKTDALQITKGEVS 8957
Qy 6838 ASERSTEAPAPASPEGAPPAACVPRHVSIRSLFYHQGESPEHGALAPGRRRHPAR 6897
Db 8958 RSEAKTELSIPKAVREDKGTVTVTASNRLGVSFRNVHVVDYR-----PSPPRNLAV 9009
Qy 6898 RRLHLKGGYIAGALP--GLEPLMEHRLVLEEAREEOA-----TLAK-----A 6940
Db 9010 TDIKAESCYLTWDAPLDNGSE--ITHYVIDKRDASRKKAEBEVTNTAVKRYGIWKLI 9067
Qy 6941 PSFETALRLPASGTHLAPGHSLSLEHDSPTSPRSSEACGEAQLPSAPSGAPTRDMGH 7000
Db 9068 PNGQYEFVRVAVNKY--GISECKSDKVVIQDP-----YRLPGPP--GKP-KVLAR 9113
Qy 7001 PQGSKQLPST-----GHPGTAQ--PERSPDSP-WGQAPAPCHPKQG-----SAPO--E 7045
Db 9114 TKGSMVSWTTPLDNGSGSPITGYWLEKREESPYWSRVSRAPITKVGLKGVEFNVPRILLE 9173
Qy 7046 GCS-PHPAVACPPGSPPGSKCAEAPLVSP-FLGQOAPAPAKASPPPLDSXMGPGDI 7103
Db 9174 GVKYQFRAMAINAAGIGPSE-----PSDPEVAGDPIFPPGP-PSCEPVKDKT-KSSI 9224
Qy 7104 SLPRGPRPGCSPGSAASQSSOVSSLRVSSOVGTGEPGSLDAEGW--TQEAEDLSDS 7161
Db 9225 SLGKHP--PAKDG-----SPIKGYIVEMQEGT-----TWKRVNEPDKLIT 9266
Qy 7162 ----TPTLQRPQEQATWRKFLSGRGGYAGVAGYGTTFAGDAGGMLGQGPMMARIAWAV 7217
Db 9267 CECVVPNLKE-----LRKYRF-----RVR-AV 9287
Qy 7218 SQSEEEQEEARASQSEEQEAEASPLPOVSARPVPEVG-----RATRS 7264
Db 9288 NEAGESFSDTGTGIPATDIOEE-----PEVFI-----DIGAQCPLVKAGSQIRIPAVI 9337
Qy 7265 SPEPTP---WEDIQGVSLVQ---IRLDSGDAEAADTISLDISEVDPAVLNLSLDYDI--K 7316
Db 9338 KGRPTPKSWEFDGKAKKAMKMGVHDIPEDAQLTAENSSVIIPECKRSHTKYSITAK 9397

Qy 7317 YLPPEFMI FRKVPKSAQPEPPSPMAEBELAEFP--EPTWMPFGEIG----- 7360
Db 9398 NKAQKQKANCRCVKMDVPGPPKDLKVSIDITRGSCRLSKWMPDDDDGDRIKGVYIEKRTID 9457
Qy 7361 -----PHAGL-----EITESEDDVALLABAAVGRKRKWSPPSRSLFHFGRHLPL 7406
Db 9458 GKAWTKVNPDCGSTTFFVVPOLLSSQYFFRVRARERFGIPPVETIQRTTARDP--IYPP 9515
Qy 7407 DEPAELGRERVKASV-----EHSIRLKGPRGEGLEK----- 7440
Db 9516 DPPIKLKILGITKNTVHLWSKPKKNDGSPVTHYIVCLAWDPTGTTKBAWRQCNKRDVE 9575
Qy 7441 -----PPRKPGGLASFRLSGLKSW-----DRAPTF-LRELSDETIVL 7476
Db 9576 ELQFTVEDLVEGGBYEFRKAVNAAGVSKPSATVGPCQPCQRPDMPSPSIDLKEFME--VEE 9633
Qy 7477 GQSVTLACQVSAQAPAAQATWSKOGAPLESSRVLISATLKNFOLL--TILVVVAE---D 7530
Db 9634 GTNNVIVAKIKGVFPPTLTWFKAPPKPDNKEPVLVYDTHVNKLVDVDTCTLVIPOQSRSD 9693
Qy 7531 LGVYTCVSNALGVTVTGVLKKAERPSSPCPDIG-----EYVADGVLLWVKPVEYGP 7585
Db 9694 TGLTITVANNLGTASKEMLNVLGREG-----PVGPIKFESVADQMTLSWFPKDDGG 9749
Qy 7586 ---VTVYIQ--CSLEGGSWTTLASDIFDCCYLTSLSRGGTYTFRATACVSKAGMG-PYSSP 7640
Db 9750 SKITNYVIERKANRKTWVHVSEPEKCTYIPKLGHEGVVFRIMAKNVKYGIGEPLDSE 9809
Qy 7641 SE 7642
Db 9810 PE 9811
RESULT 10
Q8WZ42 PRELIMINARY; PRT; 34350 AA.
ID Q8WZ42
AC Q8WZ42;
DT 01-MAR-2002 (T=EMBLrel. 20, Created)
DT 01-MAR-2002 (T=EMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (T=EMBLrel. 26, Last annotation update)
DE Titin.
GN Name=TTN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20309627; PubMed=10850961;
RA Freiburg A., Trombitas K., Hall W., Cazorla O., Fougereousse F.,
RA Cantner T., Kolmerer B., Witt C., Beckmann J.S., Gregorio C.C.,
RA Granzier H., Labelit S.;
RT "Series of exon-skipping events in the elastic spring region of titin
RT as the structural basis for myofibrillar elastic diversity.";
RL Circ. Res. 86:1114-1121(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21573839; PubMed=11717165;
RA Bang M.L., Cantner T., Fornoff F., Geach A.J., Gotthardt M.,
RA McNabb M., Witt C.C., Labelit D., Gregorio C.C., Granzier H.,
RA Labelit S.;
RT "The complete gene sequence of titin, expression of an unusual -700
RT kDa titin isoform and its interaction with obscurin identify a novel
RT Z-line to I-band linking system.";
RL Circ. Res. 89:1065-1072(2001).
RN [3]
RP SEQUENCE FROM N.A.
RA Cantner S.B.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ277892; CAD12456.1; -;
DR PDB; 1TNN; NMR; @=33480-33579.
DR PDB; 1TNN; NMR; @=33480-33579.

Db 4858 D-----DLIALGGQTVTLQAARVGESEPI SVTWKQGEVIREDDGKIMSFNGVAV----- 4907
QY 1497 RELVVOAQADAGEYSC-----EAGSORLSFHLHVAEPKAVFAKEQPASREVQAEAGTSA 1552
Db 4908 --LIIIPDVQISFGKYTKLAEANEAGSQTSGVELIVKEPAKIIER-----AELIQVTAGDPA 4961
QY 1553 TILCEVA-QAOETVTVYKDGKLSSSSKVRMEAVGCTRLVLVQEAQADAGEY-----SCK 1607
Db 4962 TLEYTVAGTPELKPWKYKDGRLPVLASKYRISEFKNVVAQLKPYSAELHDSGGQTTPEISNE 5021
QY 1608 AGQRLSFHLHVAEPKAVFAKEQPAHREVOABAGASATILCEVA-QAOETVTVYKDGKLL 1666
Db 5022 VGSSSCETTFTVLDRDIAPEFTKPL-RNVDVSVNGTCRLDCKIAGSLPMRVSMFKDGKEI 5080
QY 1667 SSSSKVRVEAVGCTRLVLVQQAQADAGEYSCBA-----CGQRLSFLHVAELPEQISERP 1722
Db 5081 AASDRYRIAFVGTASLEIIRVDMNDAGNFTCRATNSVGKSSGALIVQE-PPSFVTKP 5139
QY 1723 CRREPLVVKHEHEDIILTATLATPSAATVTLKDGVEIRRSKRHETASQGDTHTLVHGAO 1782
Db 5140 GSKD---VLPGSAVCLKSTFGQSTPLTIRWFKGNKELVSGGSCYITKEALESSELYLVK 5196
QY 1783 VLDSAIYSCRVG--AEGQDFPVQ--VEEVAAPKRLLEPVCBEL---GGTVTLACELS-P 1834
Db 5197 TSDSGTYTCKVSNVAGGVECSANLKVKE-PATFVEKLEP--SOLLKKGDATQLACKVTGT 5253
QY 1835 ACAEVVWRCGNTQPRVGRKFM--VAEGPVRSILTVLGLRAEDAGEVVCESRD----DHTS 1888
Db 5254 PIKIKITFANDREIRKESSKHRMSFVESTAVLRLTDVGI--EDSGYMCEABQNEAGSDHCS 5311
QY 1889 AOLTV-SVPRVVKFMFSGLSLST-----VVAEGGEATFOCVVSPSDVAVWFRDQALLQP 1940
Db 5312 SIVIVKESPYTFKEKPIEVLKEYDVMLLAEVAGTFPFE-----ITWFKONTILRS 5362
QY 1941 SKFPAISOGGASHSTISDLVLEDAGQ-----ITVEABGASSAALRVREAPVLFKKLEP 1996
Db 5363 GRKYKTFIQLHVLVSLQILKFVAADAGEYQCRVTNEVGSSISCARVTLREPP-SFIKKIES 5421
QY 1997 QTVEERSSVTLVELTRPWP-ELRWTRNATALAPGNVEIHAEAGRHVLHNVGFADRG 2055
Db 5422 TSLRGGTAAFOATLKGSLFITVTLWKDSDELTDNDNRMTPFNENNVASLYLSGIEVKHDG 5481
QY 2056 PFGCETPD-----KTQAKLTVEMRQVRLVGLQAVEAREQGTATMEVOLSHA-DVDSGWT 2110
Db 5482 KYVCAKNDAGIQRCSALLSVK-EPATITEEAVSIDVTQGDPAITQVKFSGTKETAKWF 5540
QY 2111 RDGLRPOQGPCHLAVRGMHTLTLGLRPEPSGLMVPKABG-----VHTSARLVVTEL-- 2164
Db 5541 KDGQBLTLGSKYKISVTDPTVSLIKIISTEKKDSGEYTFEVQNDVGRSSCKARINVLDLII 5600
QY 2165 PVSFSPRLQDVVTTTEKEKVTLBCEL--SRPNVDVRLKDGVELRAGKTMATAAQACRSL 2222
Db 5601 PPSFTKKLKMDSIKGSDIDLECI VAGSHP-ISIQWFKDDQISASEKYSFHDNTAPL 5659
QY 2223 TIYRCEPADQGVVYCDHAADASSAVKVGRTYTLIYRRVLAEDAGEIQFVAENAESRAQ 2282
Db 5660 EISQLEGTDSGTYTCSATNK-----AGHNQCSGH 5688
QY 2283 LRVKELPVLVRLPRLDKTAMEKHGVLBCQVRSASAVRWPKGSOELOQPKYELVSDGL 2342
Db 5689 LTVKEPPYFVEKPPSQODPNPNTRVOLKALVGTAFTMTIKWFKDNKELHSGAARSVKWDDT 5748
QY 2343 YRKLIISDVHAEDEDTYTC---DAGDVKTSAOFFVEB-----QSITIVRGLQDVTV 2390
Db 5749 STSLFLPAAKATDSGTIYICQISNDVGTATSKATLFEVKEPPQFIKPSFVLVLRNGQSTT- 5807
QY 2391 MEPAPAFECE--TSIPSVRPPKWLKG--TVLQAGG-----NVGLBQEGTWH 2434
Db 5808 -----FECQITGTPKIRVSWYLDGNEITAIQKHGISFIDGLATFQISGARVENSQTY- 5859
QY 2435 RLMLRRITCSTWTPGVHFTVGVKSRSSARLVVSDIPVVLTRPLEPKTGRLOQSVVLSCDPRP 2494

Db 5860 -----VCEARN-----DAGTASCISIELKVKEPPTFI-RELKPVBEVVKYSDVELECEVTG 5907
QY 2495 APK-AVQWYKDDTPLSPSEKFKMSLEGQMAELRILRLMPADAGVYRC-----QAGSAHST 2549
Db 5908 TTPPEVTVLKNRREIRSSKKYTLTDRVSFNLHTKCDPSDTGEYQOCIVSNEGSCSCST 5967
QY 2550 EVTVE-----AREVTVTGLQDAEATEBGMASFSCSELSHEDBEVWSLNGMPLYNDSFHE 2604
Db 5968 RVALKEPPSFIKKTENTTVLKSSATFOSTVAGPPPI-----ITWLKDDQLLDEDDNVY 6022
QY 2605 ISHKRRHTLVLSIQIADAGIVRASSLUKSTVSARLEVR-----RPVVFRLKALD 2654
Db 6023 ISFVDSVATLQIRSDVNGHSG-----RYTCAQKNESGVERCYAFLVLVQEPQAIVERAK 6075
QY 2655 DLSAEERTLALQCEVS-DPEAHVVRKDGVLQGPSDKYDFLHTAGTGLVHVHVSPPDA 2713
Db 6076 SVDTEKDPMTLECVAGTPELKVWKLKDGKQIVPSRYFSMSPENNVAASFRIQSMKQDS 6135
QY 2714 GLYXCHV-----GSBETRARVRVHDLHV--GITKELKTMVELEGESCSPECVLSHESASDP 2767
Db 6136 QOYTFKVENDFGSSSCDAYLAVLQONIPPSFTKKLTWMDKVLGSSIHMECKVS-GSLPIS 6194
QY 2768 AMWTVGGTVGSSSRFOATROGRKYILVREAAFPDADAGEVVFVRGLTS-----KASLIVR 2823
Db 6195 AQWFKDGKEISTSAKYRLVCHERSVSVLENNLEDTANYTCKVSNVAGDDACSGILLTVK 6254
QY 2824 ERPAALIKPLEDQWVAGEDVELCELRSRGT--VHMLKDKRAIKRSQKYDVVCEGTM 2880
Db 6255 BPPSFLVKPGRQAI-PDSTVEFKAILK--GTPPFKIKWFKDDVELVSGPKCFI GLEGST 6311
QY 2881 AMLVIRGASLXDAGEYTC-----EVEASKSTASLHVEBKANCFTBELTNLQVBEKG-TAVF 2935
Db 6312 SFLNLYSDASKTQOYTCYCHVTNDVSGDSCITMLLVTPPK-FVKKLEASKIVKAGDSSRL 6370
QY 2936 TCK-TEHPAATVTVWRKGLLELRASGHQHPQOEGITLRLTISALKAOSDVTYCD-----IG 2990
Db 6371 ECKIAGSPEIRVWFRNEHELPAFDKYRMTFIDSAVIMQNNLSTEDSGDICEAQNAPG 6430
QY 2991 QAQSRQALLVQGRVH-----IIEDLEDVVOEGSSATFRCRI SPANVEPVHWFIDKTPL 3045
Db 6431 STSGCTKIVIVKEPPVFSFPPIVELTKNAEV-----SLECELSGTPPFVWVYKDKRQL 6484
QY 3046 HANELNIDAQPGGYHV-LTLRQALKDSGTIYPEAGDQDASAAALRVTEK---PSVFSRE 3101
Db 6485 RSKKYKIASK--NFHTSIHLNVDTSDIGEHYCKAQNVEGSDTCVCTVKLKEPPRFVSK 6542
QY 3102 LTDAITIEGEDTLVCTSTCD-IPMCWTXD-GKTLRGSARCQLSHGRHQAQLITGATL 3159
Db 6543 LNSLTVVAGEPAELQAGIEGAQPIFVQWLKEKEEVIRESNIRITFVENVATLQFAKAP 6602
QY 3160 QDSGRYKCEA--GGACSSSIVRVHARVPVRFQEAALKDLEVLGGAATLRC-VLSSVAAPV 3215
Db 6603 ANAGYICQIKNDGMEENWATLMVLEPAVIVEKAGPMTVTVGTCITLCECKVAGTPELSV 6662
QY 3216 KWYCGNNVLPRGDKYLSURQEGAMLELVRLNRLPQDSGRYSC-----SFGDQTTSATLTVT- 3270
Db 6663 EWYKDGKLLTSSQKHKFSFYNKISSRLILSVEROADAGTYTFVQVNNVSKSSCTAVVDSD 6722
QY 3271 -ALPAQFIGIKRKEATEGATATLCELSTAP--VEWRKGSETLRDGDYCLRDQGNMC 3327
Db 6723 RAVEPSPFRLKNTGGVGLGASCITLCECKVAGSSISVAVMFHEKTKIVSGAKYQTTFSNVVC 6782
QY 3328 ELQIRGLAMVDAABYSC-----VCGEERTSASLITRPMFAHFIGRLRHOESTEGATATLRC 3383
Db 6783 TLQNSLSDSDMGNTYTCVAANVAGSDCRAVLTVQEPSP-FVKEPELVLPGKNVTFTS 6841
QY 3384 ELSKAAP--VEWRKGRSLRDRHSRLRQDGA VCELOICGLAVADAGEYSCVCGEERTSA 3441
Db 6842 VIRGTPPKVWFRGARELVKGRCNIFYEDTVAEELFNIDISQSSEYTCWVSNNAQA 6901
QY 3442 TLTVKAL--PAKFTTEGLRNEEAVEGATMLWCELSKAP--VEWRKGPENLRDGDYIL 3496
Db 6902 SCTRLFVKPAPAFKRLSDHSVEPGKSIILESTYTGTLPISVTWKKGDNITTSEKNI 6961

Qy	3497	QBETRC	ELQICGLAMADAGEYLC-----VCGERTSATLTIRALPARFIEDVKQ	3547
Db	6962	VTEKTC	ILEILNSTKRDAGQYSCIEINEAGRDVCG-----ALVSTLEPPYFVTELEPL	7015
Qy	3548	EAREGATV	LQCELNSAAP--VEMWKGSETLRGDGRYSLRQDGTCKELQIRGLAMADTG	3604
Db	7016	EAAVGDS	VSQCQV-AGTPEITVSWYKGDYTKLRTPPYRYFTNNVATLVFNKVNINDSG	7074
Qy	3605	EYSCVC	-----GOERTSAMLTVRALPIKFTEGLRNEEATEGATAVLRCELSKMAPE--	3656
Db	7075	EYTCKAENS	IGTASSKTVFRIQBRLPPSPARQLKDIEQTVGLPVTLTCLRLNGSAPIQVC	7134
Qy	3657	WKKGHETLR	GDGRHSRLQDGAARCELOTRGLVAEDAGEYLCMC-----GKERTSAMLTVRAM	3712
Db	7135	WYRDGVL	LRDDENLQTSFVDNVATLKLQTLDSHGQYSCSASNPLGTASSSARLTARE-	7193
Qy	3713	PSK---	FIGLRNEEATEGDTATLWCELSKAAP--VEMWKGHETLRGDGRHSRLQDGRSRC	3767
Db	7194	PKSPFFDI	KPVSIDVLAGEADPECHVTGAQPMRITWSKONKEIRPGGNYTTCVGNTP	7253
Qy	3768	ELQIRGLA	VNDAGYSVCV-----GQERTSATLTVRALPARPIEDVK-NQAREGATAVLQ	3822
Db	7254	HLRLKVG	KGDSGYTCQATNDVGDKMCSAQLSVKE-PPKFVKKLEASKVAKQGESIQLE	7312
Qy	3823	CELSKA--	APVEMWKGSETLRGDGRYSLRQDGTCELOIHLGLSVADTGEYSVCV-----GQ	3876
Db	7313	CKTSGSPEI	KVSWFRNDSLHESWKYMNSTFINSVALLTINEASAESDGYICEAHNGVGD	7372
Qy	3877	ERTSATLT	VRAPQVFPREPLQSLQABEGSTATLQCELS-EPTATVWMSKGLQLOANGRR	3935
Db	7373	ASCSTALT	VKAP-PVFTQKSPVGNALKGSDVILQCEISGTPPEPVEVWVKDRKQVRNSKKF	7431
Qy	3936	EPRLOGCTA	ELVDLOREDGTGYTCGSAQTSATLTVTA--APVRFLURELHQHEVDE	3992
Db	7432	KITSKH	PDTSILHILNLEASDVGRYHCKATNEVSGSDTCSVKFKEPPRFVKKLSDTSTLI	7491
Qy	3993	GGTAHL	CCELSRA-----GASVEMWKGSLQLPCCAKYQMVQDGA-----AELLV	4037
Db	7492	GDVEL-----	RAIVEGFQPISVVWLKD-----RGEVIRSENSTRISFDINIA TLQL	7538
Qy	4038	RGVEQEDAG	BYTC---DTGHTOSMASLSVRPRPKFKTLQSLQEQTGDIARLCQLSD	4093
Db	7539	GSPEANS	GKYICQIKNDAGMRCSAVLTVLEP-ARIIKEPEPMVTVTGNPPALECQVVTG	7597
Qy	4094	AESGAVVQ	WUHKEGVELHAGPKYEMRSGOATRELLIHOLEAKDTGEYAC-----VTGQOKTA	4149
Db	7598	TPE--LSAK	WPKDGRSLSADSKHHITFTNKVASLKIPCAEMSDKGLYSFEVKNVSGKSNCT	7656
Qy	4150	ASLRVTEPEV-	-TIVRGLVDAEVTADDEDFECSVEVSFRAGATGVQWCLQGLPLQSNVTEV	4207
Db	7657	VSVHVS	DRIVPPSPFIRKLKDVNAILGASVVLRCVSGSAPISVGMFODGNEIVSGPKCQS	7716
Qy	4208	AVRDGR	IHTLRLKGVTPEDAGTVSFHLGNHAS---SAQULTVRAPEVTILLEPLQDVQLSE	4263
Db	7717	SFSE-NVCT	LNLSLLESDDGTGIYTCVAANAVGSDCSAVLTQVEPSPFEQTP-DSVEVLP	7774
Qy	4264	QODASFO	CLRSLRAGQBARWALGVPLQANEMNDITVEQGTLLHLLTHKVTLEBAG--T	4320
Db	7775	GMSLTFT	SVIRGTGTPPFVKVFKGSRLELVPGESCNIISLEDFVTELELFEVQPLSSG	7834
Qy	4321	VSPHVG	CSSEAOQKVTKNTVVRGLNVEALGGEALFECQLSQPEVAHTWLLDDEPV	4380
Db	7835	VTNDGAS	ACSTHLLFVKEPATFVKRLADFVSFTGSPIVLEATVTGTPPISVSWIKOYE	7894
Qy	4381	RTSENAS	WFFENGLRHLLLLKNLR--PODSRCRVTFLAGDMVTSAFULTVRGMRL	4435
Db	7895	SOSERCSI	TWTEKSTILEILESTIEDYAOYSLCIENEGAGDICEALVSV---LEPPYFI	7950
Qy	4436	EPLKNAA	VRAGAQAARFTCTISEAVPVGEASWYINGAAVQPDSDTWTVDGSHQALLRS	4495
Db	7951	EPLHEAV	IGEPATLOCKY-DGTPEIRISWYKEHTKRLSAPA-YKMQFKNNVASIVINK	8008

Qy	4496	AQPHHAGEVTPACR	----	DAVASARLTVL	-----	GLPDP	-----	PEDAEVVA	4533
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8009	VDHSDVGE	--	YSCADNSVGA	SSAVLVI	KERKLPP	FFARKLKD	VHETLGLGFVA	8066
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	4534	HSHTVTLSWAA	PMSDGGGL	CGYRVEVKEG	ATGOWRL	CHEL	VPGE	CVVDGLAP	4593
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8067	NGSEPLQVSW	----	YKDG	-----	VLLKDDAN	LQTSFV	HNVATLQILQTD	8112
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	4594	FRVAAGVPGV	AGBPHVL	POTVRLAE	PPKPPV	--	QPSAPESR	QVAAGEDVSLE	4652
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8113	YNCASNP	LTAS	-----	SSAKLIL	SEHEVPP	FDLKPVS	VDLALGESGTFK	8167
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	4653	-EVIWHK	MERIO	PGGRFE	VVSOG	ROMLVI	KGFTAD	OGEEYHC	4708
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8168	IKITWAKDN	REIRPGG	NYKMTL	VENTATL	TVLVK	VGKGDAG	QYTCYASNI	8224
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	4709	FOVALSPAS	VDEAPQ	PSLPP	PEAAQEG	DLHLL	WEALARK	RMSEPTL	4768
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8225	-----	HLGVE	PPR	-----	FIKKLEPS	RIVQDE	BFTRYECKIGS	8259
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	4769	QRLPQEA	EEVAP	DSE	--	GYSTADE	LARTGD	ADLSHTSS	4827
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8260	-----	PEIKVL	WYKDE	TEIQESS	KFRMS	FVDS	-----	8291
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	4828	TSPLASK	VGAPAPS	VKPOQ	QBPLA	VRPL	GDLS	TKGLDPS	4885
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8292	-----	MENL	SVED	SGDYTC	EAHNA	AGSAS	SS	8317
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	4886	YKYRKEM	QOEG	PMFSHT	FGDT	EAQVGD	ALRL	LEC	4945
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8318	-----	TSUKV	KEPI	FRKKPH	PIETL	KGADV	HLECE	8373
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	4946	IDQLG	DC	TCL	L	LAG	DRAD	AGCYTC	5004
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8374	I-MSE	NFLTSI	HLN	DAAD	IGEY	QCKAT	NDVGS	8426
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	5005	DAPFRA	ARRL	RLR	FR	TKSP	AEV	SDEELF	5064
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8427	DISTV	VGKEV	-	QLOTT	IEGA	EPISV	VWVF	8481
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	5065	LTEAR	QAV	-----	TRFOE	MAT	IGIV	EIKLV	5116
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8482	VEPAN	AGKY	TCQIK	NDAG	MOCE	FATLS	VEPAT	8522
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	5117	PSLLT	DAAP	VP	LT	QNO	SVQ	DGYP	5175
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8523	-----	KVTG	DTCT	LECT	VAGT	PELST	TKWFK	8563
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	5176	EDQOG	HQLI	ITAV	PAD	MGVY	RCL	AE	5235
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8564	FNK	VSG	--	LKI	IN	VPS	D	8602
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	5236	PSAQY	LS	RE	QEG	TST	DB	GG	5295
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8603	-----	SDRT	VPP	PF	TRKL	KET	NGL	8643
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Qy	5296	KDGO	PLT	SA	H	IRMT	GK	L	5355
		: : : :	: : :	: : :	: : :	: : :	: : :	: : :	
Db	8644	HEG	NE</						

Db	8850	AI--LEIPSTVEDAGQYNCYIENASGK--DSCSAQILILEPPYFVKQLEBPVKSVDGSA	8905
Qy	5565	EDSKTSPASPRHGRS-----RPSS-----IOES-----SSESE	5593
Db	8906	SLOCQAGTBEIGVSWYKGDTPKLRPTTYKMHFRNNVATLVFNQVDINDSGEYICKAENS	8965
Qy	5594	DGARGEIFDIYVTYADYLPGLAEQDAITLREGQVVEV-----LDAHPL--RMLVTKP	5646
Db	8966	VGVSASTF--LTVQBQKLPFSRQLRDVGTGLPVVFDCAISGSEPIVSWSYKDGKP	9023
Qy	5647	TKSSPRQGWSPAYLDRRLKLSPEWGAABEPEFCEAVSDEYKARLSSVIOELLSSQ	5706
Db	9024	LKDSN-----VQTSFLDNTATLN-----IFKTRSLAGQYSCATNPIGSASSAR	9070
Qy	5707	AFVEELQFQSHLOHLERCPHVPIAVAGQKAVI-----FRN	5743
Db	9071	LILTE-----GKNPPFPDIRLAPVDVAVGESADFECHVTGTQPIKVSNAKDSRE	9119
Qy	5744	VRDIGRHSFLQE-----LQOCDTDDVAMCFINKQAAPFOYLEFLVGRVQAESVVVST	5798
Db	9120	IRSGGKQIYLSLNSAHLTVLKVDKGDS-----GOYTCAVNEVGKDSCTAQL	9167
Qy	5799	AIQEFYKYABEALLAGDSPQPPPLQHYLEQPVERVORYQALLKELIRKARNQONCA	5858
Db	9168	NIKE-----RLIPPSFTKRLSEVET-----	9189
Qy	5859	LLEQAVVVSALPORAENKLVLSMENVPGTLEALGEBPIROGHFIWEGAPGAMP-----	5914
Db	9190	-----FLFRNLHVICPRRDSKRTDVSVFRNMWKLSSIDLNDQVEGDDR	5967
Qy	5915	WKGNHRV-----FLFRNLHVICPRRDSKRTDVSVFRNMWKLSSIDLNDQVEGDDR	5967
Db	9210	WYKNNIEIOPTNCEITFKNNTLVLOVRKAGNNDAGLYT-----CKVS-----ND-----	9254
Qy	5968	AFEWQERDSVRKYLLOARTAIKKSWSV--KEICIGIQRALPVMRPDPFEELADCTAE	6026
Db	9255	-----AGSALCTSSYVIKE-----PKKPVFPDQHLTPVTVS	9285
Qy	6027	LGETVKLACRVGTGPKPVISWYKGVKQAVQVDPHHLITIEDPDGSCALLIDSLTGV-----	6080
Db	9286	EGEYVQLSCHVOGSEPIRLOWLKAGREIK-----PSDRCSFSFASGTAVLELRDV	9335
Qy	6081	-----DSQYMCFAASAGNSTLGKTLVQVPR-----FVNKVRASPVEG	6122
Db	9336	AKADSGDYVCKASNVAGSDTTKSKVTIKDKPAVAPATKKAADVGRLEFVSEPOSIRVEK	9395
Qy	6123	EDAQFTCTEGAPYQIRWYKDGALLTGNKFTLSE-----PRSGLLVAVIRAASKE	6175
Db	9396	TATFTAKVGGDPIFNVKX-----TKGWRQLNQGRVFIHQKGDEAKLEIRDTTK	9447
Qy	6176	DGLGYECELVNRLGARSASAEIRI-----QSPMLQ-----	6205
Db	9448	DSGLYRCVAFNEHGESNVNLQVDERKKQEKIEGDLRLAMLKKTPIILKKGAGEEEDIM	9507
Qy	6206	-----AQOCHREQLVAVEDTTTLERAQO	6229
Db	9508	ELLKNVDPKXEYKARMYGITDFRGLQAFELLKQSOEEETHRELEIERSERDEK3FE	9567
Qy	6230	EVTSLVKRLKLLGKAPGSPGTGDLTGPGPCP-----RGAPALQ--	6265
Db	9568	ELVSFIQRLSQTEPVTILKDIENQTVLKNDADVFEIDIKINYPEIKLSWYKTEKLEPS	9627
Qy	6266	-----FTGSPQPVTVG-----TSRAPAVPRVPOLLHEGP	6295
Db	9628	DKFEISIDGRHRTLKVNCKLQDQGNVRLVCGPHIASAKLTVIEPAPERHLQDVTLKEG-	9686
Qy	6296	EQEPEAIARAQSWTYP--JRMGGAW-----PGAGTGELLWDVHSHV---VRETTOR	6342
Db	9687	-----QTCMTWCQSPVNVKSE---WFRNGRILKQGRHKTEVHKHUKLTIADVRAEDOG	9739
Qy	6343	TYT--YQADITHT-----ARPPSHQVITIEDVQAOGTGTAFOFAEIIIEGDDPOPSVTVYKDSV	6395
Db	9740	QYTCKYEDLETSABELRIEABPIQTKRIQNIIVVSHQSAATFECEVSFD--DAIVTYWYKGP	9799
Qy	6396	QLVDSTRLSSQOEGTYSLSLRHVASKDAGVYTCLAO--NTGGQVLCKAEL-----LV	6446
Db	9799	ELTESQKNYFNDRGCHYMTIHNVTDPDEGVYSIARLEPRGEARSTAEILYLTKEIKLE	9858
Qy	6447	LGDNNEPDS-----BKQSHRKLHSFYFVKEE	6473
Db	9859	LKPPDIPDSRVPIPTMPIRAVPPEEIPPVVAPPIPLLLPTPEBKDPKKRI-----EVTK-	9913
Qy	6474	IGRGVGFGRVQVHKGKILCAAKFIPLRSTRQAAYRERDILAAASHPLVTLGLLDQFET	6533
Db	9914	-----KAVKDAKVVAKPKEMTPREEIVKKPPPTTLIPAKAPII-----DVSS	9959
Qy	6534	RKTULILELCSSEL---LDLRYRGVWTEAEVYVIOOLVEGLHYLHSHGLHLDIKP	6590
Db	9960	KAEFKVIMTITRKKEVQKEAEVYKQAVHKEKRVFIESFEYPD-----ELEVEP	10011
Qy	6591	SNILMVHP-----AREDIKI-----CDFGFAQNTIPAELOFSQYGSPE---	6628
Db	10012	YTEPFGQPYEEPDDEYIEIKVAKKEVHEEEDFEEGOYVEREEGYDE--GESEWBA	10070
Qy	6629	FVSPEIIQ-QNPSSEAS-----DIWANGVISYLSLTCSSPFAGESDR	6669
Db	10071	YQEREVIQVQKEVYESHERKVPKAPVEKAPPPKVKPKVLEKLEKT--SRMEEEKVQ	10129
Qy	6670	ATLNLVLEGRVSWSPMAHLSDEADKDFIKATLQAPQAPRASAQCLSHPWFLKMPABE	6729
Db	10130	VTKPYEVSCKIIVPOKPSRTVQBEVIEV-----KVPVHTTKMWISEEKMFFASHTEE	10183
Qy	6730	AHFTNTKQLKFLLARSWQRSLMSYKSIILWMSIPELLRGPPDPSLGVARHLCDRTGGS	6789
Db	10184	VS-VTVFEV-----QKEIVTEKIHV--AVSKRVEPPPKVPEL-----	10218
Qy	6790	SSSSSDNELAPARAKSL--PPSPVTHSPLLHPRGFLRPSASLPEEAEASRSTEAP-A	6847
Db	10219	--PKPAPAEVAPVPIPKVPEPPAPKPEV-----PKKPVPEKKPVVPKKEPAA	10267
Qy	6848	PPASPE-----GAGPPAAQCQVPRHSVI--RSLFVHOAGESPEHGA--	6886
Db	10268	PKPVPEVPKPVPEEKIPVPVAKKKAAPKAPVPEQKRVVTEKTIIVTQREESPPPAVP	10327
Qy	6887	-----LAGRRRHPRARRHLLKGGYIAGALPGLREPLMEHRV-----LE	6925
Db	10328	EIPKKKVPERKVPKPEEVEVPPPKVPALP--KKVPPEKVAVPVPVAKKAPPPRAEVS	10385
Qy	6926	EEAAREEQATLLAKAPSEFETALRPLASGTHLAPGH--SHSLEHSDPS-----	6970
Db	10386	KKTVVEEKRVAEKLUSFAVQVRVTRHEVSAABEWSYSEEEGVSISVYREEREBEE	10445
Qy	6971	-----TPRPSBACGE-----AQLRPSAPSGAPIRDMGHPGQSG--OLPSTG	7011
Db	10446	EAEVTEYEVNMEEPYEVVVEEKLHISKREAEAP-----EVTERQEKVILKPIPAKI	10499
Qy	7012	GHPGTAQ--PRRPSDPSWQO--PAPFCHPKQGSAPQEGGSPHFAVAPCPGSPPPGSGCK-E	7068
Db	10500	EEPPPAKVPEAPKKIIVPEKKVPAPV--PKKEKVPVPKVPPEEK-KPVEKKVPPKVIOME	10556
Qy	7069	APL-----VPSPPFLGQPOAPPAPAKASPPDLDSKMG--PG	


```
Qy 7287 GDAADTISLDISEVDPAYLNLSLDYDKYLPFPFMI FRKVPKSAQPEPPSPMAEBELA 7346
Db 10755 -----IPKLLKPPPK-VPEPPKK 10772
Qy 7347 EFPEPTWPGELGHAGLEITE-ESEDV-----DALLAEAAVGRKRKWSGSPSR 7394
Db 10773 VFEE-----KRISITKKEQVTPAAKVPKRVVAEEKVPVPRKEVAP-- 10819
Qy 7395 SLFHPGR--HLPLD-BPAELGLRERVKASVEHISRIL-----7429
Db 10820 -----PVRVPEVPKLEPEEVAPEEV---VTHVEEYLVEEEVEIHEEEEFITEEVVP 10871
Qy 7430 -----KGRPEGLEKEGPPRKP 7446
Db 10872 VIPKVPVPRKVPPEEKVPVPVPPKKKEAPPKVP 10906

RESULT 11
Q8NH4 PRELIMINARY; PRT; 658 AA.
AC Q8NH4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Obscurin (Fragment).
GN Name=OBSCN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gautel M.S.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ314904; CAC85751.1; -.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR Pfam; PF00047; I9; 1.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS00835; IG LIKE; 4.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 658 AA; 72830 MW; CB37B453FCD4AE4A CRC64;

Query Match 8.2%; Score 3389; DB 2; Length 658;
Best Local Similarity 100.0%; Pred. No. 1.3e-101;
Matches 658; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2075 RQVLRVGLQAVEAREOGTATMEVOLSHADVDSWTRDGLRFQOGPTCHLAVRGMHTLT 2134
Db 1 RQVLRVGLQAVEAREOGTATMEVOLSHADVDSWTRDGLRFQOGPTCHLAVRGMHTLT 60
Qy 2135 LSGLRPDSGLMVFKAEGVHTSARLVTELVPVSFRLQDVVTTTEKEKVTLECELSRPNV 2194
Db 61 LSGLRPDSGLMVFKAEGVHTSARLVTELVPVSFRLQDVVTTTEKEKVTLECELSRPNV 120
Qy 2195 DVRWLKDGVELRAGKTWAIQAQACRSRTIIVRCFADQGVVYCDADHAQSSASVKVQGR 2254
Db 121 DVRWLKDGVELRAGKTWAIQAQACRSRTIIVRCFADQGVVYCDADHAQSSASVKVQGR 180
Qy 2255 YTLIYRVLAEADAGEIQFVAENAESRAQLRVKELPVTILVRDKIAMKRGVLECOVS 2314
Db 181 YTLIYRVLAEADAGEIQFVAENAESRAQLRVKELPVTILVRDKIAMKRGVLECOVS 240
Qy 2315 RASQVRFWFKGSQELQPGPKYELVDGLYRKLIIISDVHAEDEDTYTCADGVKTSQAQFFV 2374
Db 241 RASQVRFWFKGSQELQPGPKYELVDGLYRKLIIISDVHAEDEDTYTCADGVKTSQAQFFV 300
```

```
Qy 2375 EQSITIVRGLQDVVMEPAPAFECETSI PSVPPKWLICKTVLQAGNVGLQEQTIVH 2434
Db 301 EQSITIVRGLQDVVMEPAPAFECETSI PSVPPKWLICKTVLQAGNVGLQEQTIVH 360
Qy 2435 RLMRLRRTCSMTGTVHFTVGKSRSSARLVVSDIPVILTRPLEPKTGRQLQSVLSCDPRP 2494
Db 361 RLMRLRRTCSMTGTVHFTVGKSRSSARLVVSDIPVILTRPLEPKTGRQLQSVLSCDPRP 420
Qy 2495 APKAVQVYKDDTPLSPSEKFKMSLEGQMAELRILRLMPADAGVYVRCQAGSAHSSTEVTV 2554
Db 421 APKAVQVYKDDTPLSPSEKFKMSLEGQMAELRILRLMPADAGVYVRCQAGSAHSSTEVTV 480
Qy 2555 AREVTVTGPLQDABATEEGHSCWASFCSELSHEDVEVNSLNGMPLYNDSFHEISHKGRHTL 2614
Db 481 AREVTVTGPLQDABATEEGHSCWASFCSELSHEDVEVNSLNGMPLYNDSFHEISHKGRHTL 540
Qy 2615 VLKSIQADAGIVRASSLKVSTARSLEVRVVKPVVFLKALDDLSAEERGTALALCEVSDPE 2674
Db 541 VLKSIQADAGIVRASSLKVSTARSLEVRVVKPVVFLKALDDLSAEERGTALALCEVSDPE 600
Qy 2675 AHVVVRKDGVLQGPSDKYDFLHTAGTGLVVDVSPEDAGLYTCHVGSEETRARVRVH 2732
Db 601 AHVVVRKDGVLQGPSDKYDFLHTAGTGLVVDVSPEDAGLYTCHVGSEETRARVRVH 658

RESULT 12
Q10465 PRELIMINARY; PRT; 7962 AA.
AC Q10465;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Elastic titin (Fragment).
GN Name=titin;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=Berkeley; TISSUE=Skeletal muscle;
RX MEDLINE=96026330; PubMed=7569978;
RA Labeit S.; Kolmerer B.;
RT "Titins, giant proteins in charge of muscle ultrastructure and elasticity."
RL Science 270:293-296(1995).
DR EMBL; X90569; CAA62189.1; -.
DR PIR; I38346; I38346.
DR HSP; P13362; IEBT.
DR GO; GO:0030017; C:sarcomere; TAS.
DR GO; GO:0008307; F:structural constituent of muscle; TAS.
DR GO; GO:0006941; P:striated muscle contraction; TAS.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR InterPro; IPR004168; PPAK_motif.
DR Pfam; PF02818; PPAK; 53.
DR SMART; SM00408; IGC2; 43.
DR PROSITE; PS00835; IG LIKE; 58.
FT NON TER 1
FT NON TER 1
SQ SEQUENCE 7962 AA; 803018 MW; B85240533CBADE58 CRC64;

Query Match 8.0%; Score 3313; DB 2; Length 7962;
Best Local Similarity 21.5%; Pred. No. 7.8e-98;
Matches 1710; Conservative 1105; Mismatches 3287; Indels 1868; Gaps 306;

Qy 10 PRFTRPKAFVSVGKDATLSQIVGNPTQVSWKEQQQPVTAGARFRLAQDGLYRLTI 69
Db 10 PAITPLQDVTSEGPQPARFCRVSGTDL-KVSWYSKDKIKIPSRFRFMTQFDITYQLEI 68
Qy 70 LDLALGDSGVYCCARNAIGEAFAAVGLQVDAEACAEQ-----APHFLLRPT 117
Db 69 ABAYPEDEGTYTFVANNVGVQSVSTANLSLEAPESILHERIEQBIEMEMKAAPVKKRIE 128
```


Db	1995	PSFLVKGROQAIPDSTVEFKAILK---GTPPPKIKWFKDDVELVSPKCFIGLEGSTSF	2051
Qy	2222	LTIYRCBFADQGVVVCDAHDAQSSASVKVQGRVTTLIYRRVLAEDAGEIQFVAENAESRA	2281
Db	2052	LNLYSDASKTGQYTCH-----VINDVG-----SDSCIT	2080
Qy	2282	QLRVKELPVTLLRPLRLDKIAEKHRGVLECOVSRASQVR--WPKGSQELQOPGPKYELVS	2339
Db	2081	MLLVTEPPKFKVLEASKIKVAGDSRSLECKIA-GSPEIRVWVFRNEHELPAQDKYRMT-	2138
Qy	2340	DGLYRKLIISDVHAEDEDTYCDAGDVKTSQAQFVEQSITIVRGLODVVMPEPAWFE	2399
Db	2139	-----FIDSVAVIQMNKLTSGD-----FI	2160
Qy	2400	CETSI PSVRPPKMLLGKTVLQAGGNVLEQEGTVHRLMLRRTCTMTGPPVHFTVGKSRSS	2459
Db	2161	CEAQNPA-----GST-----SGT-----	2174
Qy	2460	ARLVVSDIPVVLTRPLEPKTGRELOSVLSCDRPAPK-AVQWYKODTPLSPSEKFKMSL	2518
Db	2175	-KVIVKEPPVFSSPPPIVETLKNAE-VSLECELSGTPPEPVWYKDKRQLRSSKKYKIAS	2232
Qy	2519	EGQMAELRIILMPADAGVVRQQA-GSAHSTEV-TVEARE-----VTVTGPLQ	2565
Db	2233	KNFHTSIHILNVDTSDTGEYHCKAQNEVGSDTCTVTKLEPPRFVSKLNSLTWAGEPA	2292
Qy	2566	DAEATEBG-----WASFCELSHEDBEVEWSLNGMPLYNDSPHEISHKGRHRTLVLKS	2618
Db	2293	ELQASIEGAQPIFVQWLKEBEVIRSENI-----RITFVENVATIQFAK	2337
Qy	2619	IQRADAG-----IVRASSLKYSTSARLEVRVKPVVFLKALDLSABERGTLALQCEYS-D	2672
Db	2338	AEPANACKYICQIKNDGGM-R-ENMATLWV-LEPAVIVEKAGPMVTVGETCTLECKVAGT	2395
Qy	2673	PEAHVVRKQGVQLGPSDKYDFLHTAGTGLVHVDVSPEDAGLYTC-----HVGSBETRAR	2728
Db	2396	PELSVEVYKDGKLLTSSQKHFKFPYNKIISLRILSVERQDAGTYTFQOVNNGVSKSCTAV	2455
Qy	2729	VRVHDLHV--GITKRLKTMELVEGESCSFSCVLSHESASDPAMVTGGTKVSGSSRFOAT	2786
Db	2456	VDUSDRAPSPFTRRLKNTGGVLGASCLICKVAGSSPI SVA-WFHEKTKIVSGAKYQTT	2514
Qy	2787	RQGRKYTLVREAPS DAGE---VVFVSRGLTSKASLI VREPPAAI IKPLEDOWAPGED	2843
Db	2515	FSDNVCTQLNSLSDSDMGNYTCVAANVAGSDECAVLTVQEBPSPFVKPEPELVLPGNV	2574
Qy	2844	VELRCELSRAGTP---VHMLKDKRAIKRSOKYDVVCEGTWAMLVIRGASLKDAGEVTCV	2900
Db	2575	VTFTSVI--RGTPPFKNWFRGARELVKGDRCNIYFEDTVAEELFNIDISQSGEYTCV	2632
Qy	2901	EASKSTAS---LHVEBKAFCFTEELTNLOVEE-KGTAVTCTKTEHPAAVTVWRKGLLEL	2955
Db	2633	SNNAQOQCTTRFLVFKEPA-AFLKRUSDHVEPEPKSIILESTYGTGLPTISVTKWKGDFNI	2691
Qy	2956	RASGKHQPSQEGTLRLTISALEKADSDTYTCIDIGQASQAQLLVQGRRVHIITE-----	3009
Db	2692	TTSEKCNIVTEKTCILEILNSTKGDAGQYSCFI---ENAGRDVCGALVSTLEPPFYVT	2748
Qy	3010	DLEDVDVQEGSSATFRICRISPANYEPVHWFLDKTPPLHANE-----LNEIDAQPGGVHVL	3063
Db	2749	ELEPLEAAVGDSVSLQCVQVAGTPEITVSWYKGDTKLRPTPEYRFTYFNNV-----ATL	2801
Qy	3064	TLQLALKLKGSTIYFEA---GDOORASAAARVTEK--PSVFSRELTDTATITEGEDTLVNC	3117
Db	2802	VFNKVNINDSGEYTCKAENSIGTASSKTVFRIQERQLPSPFAHQDKIEQTGVLPTVLT	2861
Qy	3118	E-TSTCDIPMCWTXKDGKTLRGARSARCQLSHGCHRAQLLITCATIQDQSGRYKCEAG---GAC	3173
Db	2862	RLNGSAPIQCVWRDGVLLRDHENLQTSFVDNVATLKILOTDLSHSGOYSCSASNPLGTA	2921
Qy	3174	SSSIVRVHARPVR-----FOBALKDLEYLEGGAATLRCVLSSVAAPVK--WCYGNVLRPG	3227


```
QY 5781 YLEFLVGRVQASVVVSTAIQBYFYKYAEALLAGDPSPPPPLQHYLEQPVVERVQYQ 5840
Db 181 YLEFLVGRVQASVVVSTAIQBYFYKYAEALLAGDPSPPPPLQHYLEQPVVERVQYQ 240
QY 5841 ALLKELIRNKARNRQNCALLEGAYAVSALPORAENKLHVSLMENYPGTLGALGSPPIRQ 5900
Db 241 ALLKELIRNKARNRQNCALLEGAYAVSALPORAENKLHVSLMENYPGTLGALGSPPIRQ 300
QY 5901 HFIVMEGAPGARMPKGNHRHVFELFRNLHVLICKPRDSRTDTVSVFPRMMLSSIDLND 5960
Db 301 HFIVMEGAPGARMPKGNHRHVFELFRNLHVLICKPRDSRTDTVSVFPRMMLSSIDLND 360
QY 5961 QVEGDDRAFEVWQEREDSVRKYLLOARTAIKSSWVKEICGQIQRALPVMRPPDFEEL 6020
Db 361 QVEGDDRAFEVWQEREDSVRKYLLOARTAIKSSWVKEICGQIQRALPVMRPPDFEEL 420
QY 6021 ADCTAELGETVKLACRVGTGPKFVLSWYKOGKAVQVDPHHILIEDPDGSCALILDSLTGV 6080
Db 421 ADCTAELGETVKLACRVGTGPKFVLSWYKOGKAVQVDPHHILIEDPDGSCALILDSLTGV 480
QY 6081 DSGQVWCFAAASAGNCSTLGLKILVQVPRFVNKVRASPFVGEDAQFTCTIEGAPYPOIR 6140
Db 481 DSGQVWCFAAASAGNCSTLGLKILVQVPRFVNKVRASPFVGEDAQFTCTIEGAPYPOIR 540
QY 6141 WYKOGALLTGNKFTLSEPRSGLLVLVIRAASKEDLGLYECELVNRGLSARASAEALRIQ 6200
Db 541 WYKOGALLTGNKFTLSEPRSGLLVLVIRAASKEDLGLYECELVNRGLSARASAEALRIQ 600
QY 6201 SPMLQAQOCHREQLVAADVDTTLR-----ADQEVTSVLKLLGPKAPGSPSTGDLT 6252
Db 601 SPMLQAQOCHREQLVAADVTEQETKVPKKTIVIEETITTVKVSQRQR--SPSKSPSR 659
QY 6253 GPGQPCFRG-----APAL-----QETGSP--PVTGTSEAPVAPRVP----- 6286
Db 660 SPFRCSASPLRPLGLLAPOLLVLPNGQRRPEABPGQPVVPTLVYVTEAEHSPALPGLS 719
QY 6287 -PQPLLHEGPE-----QBPEAIARAQEWTVFIRMEGAAW--PGAGTG 6325
Db 720 GPQPKWVEETIEVVRVKMGPGQVSPTE--VPRSSSGHLFTLPGATPG 767

RESULT 14
Q72120
AC Q72120 PRELIMINARY; PRT; 8081 AA.
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE Uncoordinated protein 89, isoform b.
GN Name=unc-89; ORFNames=C09D1.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RG WormBase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
RL investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
PA Du Z., Le T.T., Wilson R.;
RT "The sequence of C. elegans cosmid C09D1.";
RL Submitted (MAY-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
```

```
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Wilson R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG WormBase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
CC -I- SIMILARITY: Contains 1 PH domain.
DR EMBL; AF003131; AAP68958.1; -.
DR HSSP; O01761; 1FHO.
DR WormBase; WEGene00006820; C09D1.1.
DR WormPep; C09D1.1b; CE34251.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000577; FGGY_kin.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR007110; Ig_c2.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001849; PH.
DR InterPro; IPR011036; PH_related.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR007850; RCDSD.
DR InterPro; IPR000219; RhoGEF.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR011511; SH3_2.
DR Pfam; PF00041; fn3_2.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF00069; Pkinase; 2.
DR Pfam; PF05177; RCDSD; 5.
DR Pfam; PF00621; RhoGEF; 1.
DR Pfam; PF07653; SH3_2; 1.
DR ProDom; PD000001; Prot_kinase; 2.
DR SMART; SM00060; FN3; 2.
DR SMART; SM00408; IGC2; 24.
DR SMART; SM00233; PH; 1.
DR SMART; SM00325; RhoGEF; 1.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS00010; DH_2; 1.
DR PROSITE; PS00933; FGGY_KINASES_1; UNKNOWN_1.
DR PROSITE; PS00853; FN3; 2.
DR PROSITE; PS00835; IG_LIKE; 50.
DR PROSITE; PS00003; PH_DOMAIN; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 2.
DR PROSITE; PS00002; SH3; 1.
SQ SEQUENCE 8081 AA; 894245 MW; 67C804953CF62228 CRC64;

Query Match 7.7%; Score 3182.5; DB 2; Length 8081;
Best Local Similarity 21.1%; Pred. No. 1.3e-93;
Matches 1891; Conservative 1190; Mismatches 3350; Indels 2549; Gaps 375;

QY 27 ATLSLCIVGNPTPOVSWKXQDPVTAGARPLAQDGD-LYRLTILDLALGDSQYVCRAR 85
Db 565 AKMCLKVTGYPLDITWYKDDVQLHDEDRHTFYSDEDDGFAMTIDPVQVTDTRGYTCMAT 624
QY 86 NAIGEAFAAV--GLQVDAEACAEQAPHLRPTTSIRVRGESEATFRCRVGSPPRPAVS 142
Db 625 NEYGQASTSAFFRVLKVEKEA----PPAFVTKLRDKEGQDVDFEVEGEVWPFPELV 690
QY 143 WSKDGRRLGEPDGPVRVVEELGEASALRTRARPRDGGTYEVRANPLGAASAAALVVD 202
Db 681 WLVDQDLRPSHDFRLQYD--GQTAKLEIRDAQDDTGVYTVKIQNEFGSIESKAELEFVQ 738
QY 203 SDAADTASRPGTSTAALLAHQRRREARAEAGAPASPPSTGTCTCTVTEGKHARLSCYVT 262
Db 739 ADPKKHVAP-----EFQATIEVECD-----EGEEVFRKSVIT 772
```


Db 2504 ITSDGHYKIVBEDGSLKLSIQTCLBQKGRITTIKAESSEFGVABCSASLGV-----VK 2556
QY 2260 RRVLAEDAGEIOFVAENAESRAQLRVKELPVLTVRPLRDKIAMEKXRGVLECOV---SRA 2316
Db 2557 GRPMAKPA-----FQSDIA-----PINLT-----EGDTLECKLLITGDP 2590
QY 2317 SAQVWFVKGSLQBPQKVELV-SGLVRKLIISDVHAEDEDTYC-----DAGDVKTSAQ 2371
Db 2591 TPFVKWYIGTQIVCATEDTEISNANGVY-TMKIHGVTADMTGKIKCVAYNKAQEVSTBEG- 2648
QY 2372 PFVEQSITIVRGLQDVVMEPAPAFWE---CETSI-----PSVRPPK 2411
Db 2649 -----PLKVAPIPVEFETSJCDATCREGDTLKLRAVLLGPEPEPVV---S 2690
QY 2412 WLLGKTVLQAGNVGLEQEGTVHRLMLRRTCTMTGPHVFTV---GKSRSARLWVSDI 2467
Db 2691 WYVNGKLEESQNIKIHKSEGTYYTIIKDITCDISGVVCEAINEYGKATSEATL----- 2745
QY 2468 PVVLTRPLEP-----KTG-RELQSVVLSCDFRPAPAKAVQWKDDTPLSPSEKFK 2515
Db 2746 -LVLPGRBPPDFLEWLSNVRAARTGPKVVKVFTGDPKP---SLTWYINNKEILNSDLYT 2801
QY 2516 MSLEQMAELRILRLMP-----ADAGVYRQOA-----GSAHSSTEVTVREARE 2557
Db 2802 IVTDDKTSLTINSFNPVDVHVEIICKAENDAGEVSCANNITVYTMDFSESESEAOABE 2861
QY 2558 -----VTVTGP-----LQDAEATEEGWASFCELSHEDEB---VE 2589
Db 2862 FVGDDLTDESLREBHRTPPVMAKPFITIKUTKAKKGHSAPFECVV---PDTKGVCKK 2919
QY 2590 WSLNGMPLYNDSFHEISHKGRHT-----LVLKSTORADAG-----IVRASLKV 2634
Db 2920 WLKDGKEI-----ELIARIVQTRTGPGEHITQELVLDNVTPEDAGKYTCIVENTAGKD 2973
QY 2635 STSARLEV-----RVKPVFELKALDLSABERGTLALOCEV-SPEAHVVRKDGQV 2695
Db 2974 TCEATLVIESLEKKEKKAPEFIVALQDKTKTSEKVVLECKVIGBPKPVSWLHNKT 3033
QY 2686 LQPSDKYDFLRTAGTGLVHVDSPEDAGLYTCHV---GSEETRARVRVHDLHVGITXR 2741
Db 3034 I-TQESITVESVEGVERVITTSSELSHGQKYCTIAENTEGTSKTEAFITVQGEAPVTKK 3092
QY 2742 LKTHEVLEBEGSCFECVLSHESADPAWTVGGKT-----VGSSSRF-----QATROGRK 2791
Db 3093 LQNKELSTIGEKLVLSCVKGSPQPHVDYFYSFETTKVETKITSSSRIAIEHDQNTNTHWM 3152
QY 2792 YI-LVUREAAPS DAGEVVFVSGRLTSKASLIVRERPAALIKPLEDQWAPGEDVELRCEL 2850
Db 3153 VISQITKEDIVSYKAIATNSIGTATSTSKITTKVEAPVFEGLKKTSVKEKEEIKMEVKV 3212
QY 2851 SRAGTPVHLKDKAIRKSQKYDVV---CEGTMAMLVIRGASLKDAGBYTCVEASKSTAS 2908
Db 3213 GGSAPDVWFKDDKPFVSEDGHNKKNPETHGVFTLVVKQAATTDAGKYTAKASNPAQTAE 3272
QY 2909 LHVEKAN-----CFTEEL---TNLOVEEKGTAVFTCTKTEHPAATVWRKGLLELRASGK 2960
Db 3273 SSAAEAVTQSLKPTTFVRELVTTVKNETATLSVTYKGV-PDPSVFWLKDGPQVQTDSS 3331
QY 2961 HQPSQ-BGL-TLRLTISALEKADSDTYTC-----DIGOAQRAQLLVQGRV---HIIEBLE 3012
Db 3332 HVIKAVGSGSYSITIKDARLEDSGKYACRATNPAGEAKTEANFVAVNVLVPFVEFKLS 3391
QY 3013 DVDVQEGSSATFCRISGPANYEPVHVFLDKTPHLANELNEIDAQPG-GYHVLTLRQALQK 3071
Db 3392 PLEVKEKESTLTSVKVVGTPPEPSVEWFKDDTPTSIDNVHVIOQTAVGSPSLTINDARQ 3451
QY 3072 DSGTIYF-----EAGDQASALRVTEK--PSVFSRELTATITEGEDLLTIVCTETCDI 3124
Db 3452 DVG-IYSCRARNEAGEALTANFGIIRDSIPPEFTQKRLPLEVREBQETLTLKVTVITGTPV 3510
QY 3125 P-MCWTXDKGTLRGARSARCOLSH-----EGHRAQLLITGATLQDSGRYKC-----EAGGA 3172
Db 3511 PNVEWFKDDKPI-----NIDNSHIFAKDEGSGHHT-LTIKQARGEDGVYTCATNEAGEA 3565

QY 3173 CSSSIVRVHAR---PVRFOEALKOLEVLEGAATLRCVLSVAAP-VKWCGYNNVLRPGD 3228
Db 3566 KTTANMAVQSEIEEAPL-FVOGLKPYEVEQKPAELVVRVEGKPEPEVKNFKDGVPIALDN 3624
QY 3229 KYSLRQBG--AMLELVVRNLRPODSGRYSC-----SFGDQTTSATL-----TVTAL 3272
Db 3625 QHVIEKKGNGSHTLVIKDTNNADFGKYTCQATNKAKGDETVGELKIPKYPKSFEXQTAABE 3684
QY 3273 PAQFIGLKRKEATEGATATLRCELSATA--PVEWRKGSETLRDGDYRCLR--QDGAMCE 3328
Db 3685 KPUIFIEPLKETFAVEGDTVVLECKVKNKESHPOIKFFKNDQFVEIGQHMQLLEVLEDGNI-K 3743
QY 3329 LQIRGLAMVDAAEYSC-----VCGEERTSASLIT----- 3357
Db 3744 LTIQNAKEDVGAYRCEAVNVAGKANTNADLKQPAKVBEHVTDESGQLLEEIQOFTVVG 3803
QY 3358 -----RPMPAHFIGRLRHQESIEGATATLRCELSKAAP---VEMRK-GRESLRDG 3403
Db 3804 DTASSKTDTCRGAP-BFVELLRSCVTVEKQAILKCKV-KGEPRPKIKMTKEGKEVEMSA 3861
QY 3404 DRHSLRQDGAVCLOICGLAVADAGEYSCVCGERTSA-----TLTVKALP-----AK 3451
Db 3862 RVRAEHKDDGTLLTFDNTVQADAGEYRCEAENEYGSAMTEGPIIVTLEGAPKIDGEPD 3921
QY 3452 FTEGLRNEEAVEGATAMLWCELS-KVAP-VEMRKGPENLRDGDYILR--QEGTRCELOI 3507
Db 3922 FLQPVKPAVVTVGETAVLBGKISGKPKSPVKYKNGEELKPSDRVKIENLDDGTQ-RLTV 3980
QY 3508 CGLAMADAGBYLVCVQOE---RTSATLTLIRALPAR-----FIEDVKNQEAAREGATAVLQC 3559
Db 3981 TNAKLDDMDEYRCEASNEFGDVMSDVTLTVKE-PAQVAPGFFKELSAIQVKETETAKFEC 4039
QY 3560 ELNSAAP-VEMRKSETLRDGR--YSLRQDGTCKELOIRGLAMADTGEV---SCVCGQ 3612
Db 4040 KVSGETPDVKNFKDGTGPLKEDKRVHEESTDDGTQ-RLVIEDSKTDDQGNRIEVSNDAGV 4098
QY 3613 ERTSAMLTVRALP---IKFTEGLRNEEAETEGATAVLRCEL-SKQAPVEMWKGHETLRDGD 3668
Db 4099 ANSKVPLTV--VPSETLKIKKGLTDVNVITQGTCKILLSVEVEGPKTVKMYKGTETVTSQ 4156
QY 3669 RHLRQ-DGARCLOIRGLVAEDAGBYLWCMGKE---RTSAMLTVR-----AMPSKFI 3717
Db 4157 TTKIVQVTESEYKLEIESAEMSDTGAYRVVLSTDSFSEVSATVTVTKAAEKISLPS-FK 4215
QY 3718 EGLRNEEAETEGDTATLWCCEL-SKAAPVEMRKGHETLRDGRHSLRQDGSRCLOIRGLAV 3776
Db 4216 KGLADQSVKGTPLVLEVEIEGPKDVVKYKNGDEIKDGKVEDL--GNGKYRLTIPDFQE 4273
QY 3777 VDAGEYSCV---CGQERTSATLTVRALPARFIEDVKNQEAAREGATAVLQCEL-SKAAPV 3831
Db 4274 KDVGESYVTAANEABEIESKAKVNVSAKP-EIVSGLVPTTVKQGTATFNVKVPKGV 4332
QY 3832 EWRKGSETL-----RGDRYSLRQDGTRCLOIHLGSLVADTGEYSCV---CGQERTS 3880
Db 4333 KMYKNGKEIPDAKTDKNGDGSYS-----LEIFNAQVEDAADYKVVVSNDAAGDADS 4383
QY 3881 ATLTVRAP-----QPVPREPLQSLQABEGSTATLQCELSSEPTATVWMSKGLQLOANG 3933
Db 4384 AALTVKLADDDKDKVKVPEIVSGLIPTTVKQGTATFNVKVPKGVKVKYKNGKEI-PNA 4442
QY 3934 RREPRIQGTAEALVLQDLOREDTGEYTCGSOA-----TSATLTVTAAVPRFLRELQHOE 3989
Db 4443 KAKDNGDG-SYSLEIPNAQLDDTADYKVVVSNDAAGDADSAAALTVKLPGIAIVKGLEDAE 4501
QY 3990 VDEGGTAHLCCELSRAGASVEWRKGSLOLFPKCAKQWODG-AAEALLVRGVEQEBDAGDY 4048
Db 4502 VPKGKAVLQVETNKPKPEIKWYKNGKEITPDKAQPGSDGNKPKQLPVPDAGDDDAEY 4561
QY 4049 ----TCDTGTHT-QSMASLSVRVP--RPPEKTRLOSLEQETGDI-ARLCCOLSDAESGAVV 4100
Db 4562 KVVLTDEDGNTADSSCALTVKLPAKEPKI---IKGLEDDQVVISGSPIKLEIETSGSPKTV 4618

Db 6158 NKIVFNMPTIFDCLVVGHPAPEVWFHNGKKIVPGGR1K-IQSCGGSHALIILDTTLE 6216
QY 6176 DLGLYECELVNLRGASARAEALRIQSPMLQAOEQCHREQLVAAVEDTTLERADQEVTSVL 6235
Db 6217 DAGEYVATAKNHSGSSANVLDVTFPL-----DSIKFNGEIDVTFYL 6260
QY 6236 KRLLG-----PKAPGSTGDLTGF-----PCPRGAP----- 6262
Db 6261 TEYGFKGLNTASLTPPPDRGPFKEVTHGYLTLSMIPTRAPPRYPQVYVIEIRELPE 6320
QY 6263 ---ALQETGQPPVT-----GTSE-APAVPP-----RVPOPLIHE 6293
Db 6321 KOWSLLEYNIPEVCKVRNLEKSYQFVRAENYIGISDPSPASPPSLMAPPOVFDR 6380
QY 6294 -----GPEOEPAIARAQE-----W-----TVPIRMEG----- 6316
Db 6381 RTNKVPIPLDPAEKALDMRYSEQYACAPFPGVVEKRYCAENDTLIIVLVNSGFPDPPD 6440
QY 6317 AAWPGAGTGELLWDVHSHVVRRTTQRT-----YTQOAIIDHTVA----- 6354
Db 6441 IKWKFRG-----WDI-----DTSSPTSKCKVYTYGSGSETTLAITGFSKENVQYQCFAP 6489
QY 6355 -----RPSMOVTIEDVQAOCTGTAQ-----FEALIEGDPQSVTHYKDSV 6395
Db 6490 NDYGAQOQIMVDLATRPNFIQLVN-----KTFSSAQPMRMDVRVDGEPFPELKMKEWR 6545
QY 6396 QLVDSTRLSSQOQEGT-TYSLVLRHVASKDAGVYTCLAQNTGGQVLCKAELLVLGDNEDPD 6454
Db 6546 PIVESSRIKFVQDGPVLCSLIINDPMWRDSGIYSCVANNDAGQATTSCTVTVEABGDYND 6605
QY 6455 SEKQSHR-----RKLSHFVEVKEETGRGVFGVFKVRQVHGKNGKILCAAKFIPILRSRTRAQA 6509
Db 6606 VELPRRRVTIESRRVRELYEISE-----KDEKL--AAEGAPFRVKEKATG 6648
QY 6510 -----YREDRILAALSHPLVTGLLDQFETRKTLILILELCSSEELLDRLYR 6555
Db 6649 REFLLAQLRPIDDALMRHVDHNSLDPGIVQHRVLRDEKALVAVFD--NANSTIDGLSS 6706
QY 6556 -----KGVVTEAEVKYIIOQLVEGLHLYSHGVHLHDIKPSNLMVHVPAREDIK 6604
Db 6707 LAHPGVEIAEPKGVNRETQVRFVRLKLLKALHMDLRIALHLDLRPETILL--QDDKLK 6763
QY 6605 ICDFGAQNIPTAELOFQOYSGSPFVSEIITQONPVSEASDIWAMGVISYLSLFCSSPFA 6664
Db 6764 LADFGQARLLRGLITGIBKSGSPFVSEIIVRSYPLTLATDMWTGVLTVYLLTGLSPFH 6823
QY 6665 GESDRATLLNVLGKRVSWSSMAAHLSEDAKDFIKATLQAPQAPRPSAAQCLSHPWF--- 6721
Db 6824 GDNDETIANVDSQCFD--SSPL-GNFYSYDAGDFVKLLTEIPVSRLLVDEALDHPWINDE 6881
QY 6722 -LKSNPABEAHFINTKQLKFLARSRWORSYKSLVWRS-----IPELLRGP----- 6770
Db 6882 KLKTEP-----LSADTLREFKYQHKWLE-----RRVFVQOTPSEQILEALGPATAQAQ 6930
QY 6771 PDSFSLGVARHLR-----DTGSSSSSSSSSSSNE- 6799
Db 6931 QNAPVAPGRPAEIIYDLRIQPKPPPTVEYVFPQPKHEPFFIDFEGQLIDGDAFRPE 6990
QY 6800 ----LAPPARAKSLPPSP-----VTHS-----PILHPRGFL 6826
Db 6991 GTGFEGRHPQPPQIPQPPQRNQAAHDSRRHEQQPHQGOQPORIPVDQYGRPLVDPR-YL 7049
QY 6827 RPSASLPEEAASERSTEAPAPSPGAGPAPAAQCGCVPRHSVIRSLFYHQAGSPSEHGA 6886
Db 7050 NDPSHRPSLDDAPPYVDKYNPNVHFDFKYGRPMA-----PONLEKRLKLIPODKGTPSHS- 7104
QY 6887 LAPGSRHRPARRRHLKGGYIAGALPG-----LREPLMEHRVLEEAAR-----E 6931
Db 7105 -KKEKTHQVATPIL-----ASPGDQOQOQKI PMRMINGERREREIEEBIANRILSDIS 7155
QY 6932 EQATLLAKAPSFETALRLPASGTHLAPGHSLSLHSDSPSTPRPSEACGAEQALRPSAPSG 6991
Db 7156 EEGSIAGSLASLED-FEIP-----KDFQVEASEPSTPTLTPE----- 7191

QY 6992 GAPIRDMGHPCQSKQLPSTGCHPGCTAQPERSPDPSPMQOPAPFCHPKGSAPQEGCSHPH 7051
Db 7192 -VTIRE-----TIKPTPSPSTSPQKSPVP----- 7214
QY 7052 AVAPCPPGSPGPGSKEAPLVPSPPFLGQOQAPPAPAKAS-----PPLDSKM--G 7099
Db 7215 -----QPQGLLIIPAKVTYSDSILAGLPAADKKVLEDA 7246
QY 7100 PGDISLP-GRP--KPGPCSSPGSSASQASSQVSLRVGSSQVGTGTEPQPSLDABGWTOEAE 7156
Db 7247 ENDPSIIPVGAPFLFLEGLHSDLTDTTTSASGL--IKVTSPLAINLSPNPK-----SPRRS 7298
QY 7157 DLSDSSTPLQORPOQATMRKFSGLGRCGYAGVAGYGTAFAGGAGGMLGQGPMMARIATA 7216
Db 7299 TPGTKSPVVLSPROHEHNEVLIIATRG-----KPGFLPPGEL----- 7335
QY 7217 VSQSEEBEAEARASQSEBQOEAREASPLQVSAREPVEVGRAPTRSSPEPTWEDIGQ 7276
Db 7336 ---AEDIDDEDAFMDRKKQ----- 7352
QY 7277 VSLVQIBDLSGDAEAD-----TISL-DISEVDP--YMLSDL-----YDIKY 7317
Db 7353 ---VKPDHGDENDFKDEKERLEKDKNRRTVNLDDLKYPSPAFYKDDSDFGHPGYDIDA 7409
QY 7318 LPFE--FM1----FRKVPKSAQPEPPSPMAEEELAEPEPTWMPGSELG-PHAGLEITEE 7370
Db 7410 TPWDSHYQIQPDYTYLMAARGAANRSVRNRYREELFGMGAPTVK-QGFLGVNRNDRITVRER 7468
QY 7371 SEDVDALLAAVAVGRKKNSSPSRSLFHPFGRHLPLDEPAELGLRERVKASVREHISRLK 7430
Db 7469 RRYTD-ILRETTQGLEPKSHEQSTAL-----LQKAPSATAIERIKIEKVTPC-- 7516
QY 7431 GREGLEKEGPPPKKCLASFRLSGLKSWDRAPTFIRELSDETIVVLGQSVTLACQVSAQP 7490
Db 7517 -----ATKKNDDGTF-----APIFTARLDVLRKQNPATFCAVSASP 7555
QY 7491 AAQATWSKOGAPLESSRVLISATLKNFQLLTLLVWVAEDLVVYTCVSNALGTVTVTGV 7550
Db 7556 APKVTWDFQKILLESNDRVITIEQD-NNVARLILNHAAPYDLGEVYCTAINEYGTDKSSCR 7614
QY 7551 LRKAERSSSPCPDIOGEVAD-GVLLVWKPVESYGP-----VTYIVQCSL-----EGGSW 7599
Db 7615 LISGETPSRGRPE-AELSDSTEIFIQWEAPE--GPTYLEGITYRLEYRVAGPNHDGDPW 7671
QY 7600 TTLASIDFCCLYTLKLSRGTYTERTACVSKAGMGYS-----SPSEQV--- 7644
Db 7672 ITVSEKIDDESIVKHLSPGIYQFRVTAQNGFGLGPLSSRIVOTHGKGAPKLQIDVL 7731
QY 7645 -----LLGGPSHLASEESQGRSA-----QPLPSTKTFAFOTQ 7677
Db 7732 KSEIRLVNVMPQKSTNQLGGISE-ESEEDSEARTANEDMKNLQLOTDDPTGRFOIGGL 7790
QY 7678 IQGRFSVVRQWCEKAS-GR-A-LAAKIIIPYHPKDKTAVLREYALKGLRHPHIAQLHAAY 7735
Db 7791 KFKGRFSVIRDAVDSTTEGHAHCAVKI--RHPSSE--AISEYESLRDGGHENVORLIAAF 7846
QY 7736 LSPRHVLVILELCSGPELLPCLAEARASYSSEVKDLMQWLSATQVYLNHQLHLLHLSRSE 7795
Db 7847 NNSNFYLLSERLY-EDVFSRFRVFNDDYITEEQVALTMROQVTSALHFLHFKGTAHLVDNPH 7905
QY 7796 NMIITEYN--LLKVVLDLNAQSLQSKBVLPSKFKDYLET--MAPEL-LEGQAVPQTDI 7850
Db 7906 NIMFQSKRWVVKLVDFGRAQKVS-GAVKPD-----FDTKWASPEFHIPETPVTVQSDM 7959
QY 7851 WAGVTAPTMLSAEYVPSSEGGARDLQRLKGLVRLSRVYAGI-----SGGAVAFURSTL 7905
Db 7960 WGMGVVTFCLLAGFHPFTSE--YDREEEIKENVINV-KCDPNLIPVNASOECLSFATWAL 8016
QY 7906 CAQPMGRPCASSCLOCPMLTTEGPACSRPAPVTFPTARLR 7945
Db 8017 KKSFVRMETDEALSHKFLSSDPSMVRRESIKYSASRLR 8056

RESULT 15

Q8NH8 PRELIMINARY; PRT; 646 AA.
ID Q8NH8;
AC Q8NH8;
DT 01-OCT-2002 (Tremblrel. 22, Created)
DT 01-OCT-2002 (Tremblrel. 22, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Obscurin (Fragment).
GN Name=OBSCN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Gautel M.S.;
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ314898; CAC85746.1; -.
DR HSP; Q10466; 1BPV.
DR InterPro; IPR003961; FN III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003598; IG_c2.
DR Pfam; PF00041; fn3; 1.
DR Pfam; PF00047; ig; 3.
DR SMART; SM00060; FN3; 1.
DR SMART; SM00408; IGC2; 1.
DR PROSITE; PS50853; FN3; 1.
DR PROSITE; PS50835; IG_Like; 4.
FT NON_TER 1
FT TER 1
SQ SEQUENCE 646 AA; 70258 MW; B04F5B084323D68 CRC64;

Query Match 7.5%; Score 3113; DB 2; Length 646;

Best Local Similarity 93.5%; Pred. NO. 1e-92;

Matches 605; Conservative 16; Mismatches 24; Indels 2; Gaps 1;

QY 421 NLLRKLPRKTAVRVGDGTAMFCVELAVPGVPHWLRNQEEVAVAGGRVAISAEGRHTLTIS 480
DB 1 NLLRKLPRKTAVRVGDGTAMFCVELAVPGVPHWLRNQEEVAVAGGRVAISAEGRHTLTIS 60
QY 481 QCCLEDVGQVAFMAGDCOTSTRFCVSAAPKPPLOPPDPVVKARMESVILSWSPPHGE 540
DB 61 QCCLEDVGQVAFMAGDCOTSTRFCVSAAPKPPLOPPDPVVKARMESVILSWSPPHGE 120
QY 541 RPTVIDGYLVEKKLGTVTWIRCHEAEWVATPELTVDVABEGNFPQFRVSALNSFGQSPY 600
DB 121 RPTVIDGYLVEKKLGTVTWIRCHEAEWVATPELTVDVABEGNFPQFRVSALNSFGQSPY 180
QY 601 LEFGTWHLAPKLAVRTPKVAQVEGSEVTFSDLTIVASAGWFLDQALKASSVYEH 660
DB 181 LEFGTWHLAPKLAVRTPKVAQVEGSEVTFSDLTIVASAGWFLDQALKASSVYEH 240
QY 661 CDRTRHTLTIREVPASLHGAOLKPVANGIESSIRMEVRAAPGLTANKPPAAAAAREVLARL 720
DB 241 CDRTRHTLTIREVPASLHGAOLKPVANGIESSIRMEVRAAPGLTANKPPAAAAAREVLARL 300
QY 721 HEEAQLLAELSDQAAAVTWLKDGRTLSPGPKYEQVQASAGRRVLLVRDVARDDAGLYECVS 780
DB 301 HEEAQLLAELSDQAAAVTWLKDGRTLSPGPKYEQVQASAGRRVLLVRDVARDDAGLYECVS 360
QY 781 RGGRIAYQLSVQGLARFLHKDMAGSCVDVAVAGGPAQFECETSEAHVHVHWYKDGMLGHS 840
DB 361 RGGRIAYQLSVQGLARFLHKDMAGSCVDVAVAGGPAQFECETSEAHVHVHWYKDGMLGHS 420
QY 841 GERFLOEDVGTGRRHLRAATVTRQDSGTYSYCRVGEDSVDFRLRVSEPKVFAKEQLARRKL 900
DB 421 GERFLOEDVGTGRRHLRAATVTRQDSGTYSYCRVGEDSVDFRLRVSEPKVFAKEQLARRKL 480

QY 901 QAEAGASATLSCEVAQAQTEVTWYKDGKLSSSSKVCMETGCTRRLLVVQAGQADAGEY 960
DB 481 QAEVGASATLSCEVAQDQMEVTWYKDGKLSSSSKVHVEAVGCMRRLLVVQVQVQADSGEY 540
QY 961 SCEAGGQRLSPHLDVKEPKVVFADQVAHSEVQAEAGANATLSCEVAQAQAEVVMYKDGK 1020
DB 541 SCEARGQKVSFRLD--EPKVVFKEQLARRKLAAGASATLSCEVAQAQTEVTWYKDGK 598
QY 1021 KLSSSLKVHVEAKGCRRLRVVQAGKTDAGDYSCEARGQKVSFRLHI 1067
DB 599 KLSSSSKVCMETGCTRRLLVVQAGQADAGEYSCAAGGQRLSFHLDV 645

Search completed: March 18, 2005, 17:58:56

Job time : 861 secs

THIS PAGE BLANK (uspto)

[illegible]

REFERENCE	Unpublished
3 (bases 1 to 2534)	
Isogai,T. and Yamamoto,J.	
Direct Submission	
Submitted (04-JUL-2002) Takao Isogai, FMJ Project (HRI Team) / 2-6-7	
Kazuo-Kamatari, Kisarazu, Chiba 292-0812, Japan	
(E-mail:genomic@hri.co.jp, Tel:01-438-52-3975, Fax:01-438-52-3986)	
NEO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB) / cDNA library construction: Heiichi Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; Clone selection for full insert sequencing: HRI and RAB.	
RAB; annotation: HRI and RAB.	
location/Qualifiers	
1..2534	
/organism="Homo sapiens"	
/mol_type="mRNA"	
/db_xref="taxon:9606"	
/clone="TBS12016848"	
/feature_type="cfeature"	
/clone_id="TBS12"	
/note="Cloning vector: pME18SF13"	
ORIGIN	
Query Match	100.0%; Score 401; DB 9; Length 2534;
Best Local Similarity	100.0%; Pred. No. 7e-71;
Matches	401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1	GAAGACCACTGGTGGCGCCGACGAGACCTGAGCTGCGCTGTGACTGTCAACGGCG 60
808	GAAACCACTGGTGGCGCCGACGAGACCTGAGCTGCGCTGTGACTGTCAACGGCG 867
61	GGAAGCGCCGCTGCACTGGCTGAAGAAGCAAGGCGCATCCGCAAGACAGACAGATAT 120
868	GGAAGCGCCGCTGCACTGGCTGAAGAAGCAAGGCGCATCCGCAAGACAGACAGATAT 927
121	GTGGTCTGCGAGGACAGATGGCATGTGTCTCATCCGCGGGGCTCTCTCAAGGACCG 180
928	GTGGTCTGCGAGGACAGATGGCATGTGTCTCATCCGCGGGGCTCTCTCAAGGACCG 987
181	GGCGAGTACAGTGTGAGTGGAGGCTTCCAAAGACACAGCAGCTTCATGTGGAGAA 240
988	GGCGAGTACAGTGTGAGTGGAGGCTTCCAAAGACACAGCAGCTTCATGTGGAGAA 1047
241	AAAGCAATGCTTCCACAGAGACTGACCAATCTGCAGGTGGAGGAGAAAGGACACACT 300
1048	AAAGCAATGCTTCCACAGAGACTGACCAATCTGCAGGTGGAGGAGAAAGGACACACT 1107
301	GTGTTACAGTCAAGACAGACACCCCGGAGCCACAGTACCTGGCGCAAGGACCTTTG 360
1108	GTGTTACAGTCAAGACAGACACCCCGGAGCCACAGTACCTGGCGCAAGGACCTTTG 1167
361	GAGCTACGGGCTCAGGAGACACACAGCCAGCCAGGAGG 401
1168	GAGCTACGGGCTCAGGAGACACACAGCCAGCCAGGAGG 1208
RESULT 3	
BC043916	2761 bp mRNA linear PRI 23-JAN-2003
LOCUS	Homo sapiens, similar to titin, clone IMAGE:3845393, mRNA, partial
DEFINITION	cde.
ACCESSION	BC043916
VERSION	BC043916.1 GI:27881999
KEYWORDS	
SOURCE	Homo sapiens (human)
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE	1 (bases 1 to 2761)
AUTHORS	Strauberg,R.

Db 8751 GCGGATACACGTGTGAGGTGAGGCTTCACAGACACACCGCTCCATGTGAGAGA 8810
Qy 241 AAAGCAAACTGCTTACAGAGAGAGCTGACCAATCTGACGTGAGAGAGAGAGAGAGCT 300
Db 8811 AAAGCAAACTGCTTACAGAGAGAGCTGACCAATCTGACGTGAGAGAGAGAGAGAGCT 8870
Qy 301 GTGTTACAGTGCAG 360
Db 8871 GTGTTACAGTGCAG 8930
Qy 351 GAGCTACGGGCTCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 401
Db 8931 GAGCTACGGGCTCTCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 8971

RESULT 7
LOCUS CQ445545 1005 bp DNA linear PAT 30-JAN-2004
DEFINITION Sequence 11305 from Patent WO0192523.
ACCESSION CQ445545
VERSION CQ445545.1 GI:41417003
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Shinketo, R.A. and Leach, M.D.
TITLE Human polynucleotides and polypeptides encoded thereby
JOURNAL Patent: WO 0192523-A 11305 06-DEC-2001;
Curagen Corporation (US)
FEATURES
source 1. 1005
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 98.0%; Score 393; DB 6; Length 1005;
Best Local Similarity 98.8%; Pred. No. 3.3e-69;
Matches 396; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GAAGACCAAGTGGTGGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 358 GAAGACCAAGTGGTGGCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 417
Qy 61 GAAGACCCCTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 120
Db 418 GAAGACCCCTGCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 477
Qy 121 GTGCTGCGAG 180
Db 478 GTGCTGCGAG 537
Qy 181 GCGGAGTACACGTGTGAG 240
Db 538 GCGGAGTACACGTGTGAG 597
Qy 241 AAAGCAAACTGCTTACAG 300
Db 598 AAAGCAAACTGCTTACAG 657
Qy 301 GTGTTACAGTGCAG 360
Db 658 GTGTTACAGTGCAG 717
Qy 351 GAGCTACGGGCTCTCAGGAG 401
Db 718 GAGCTACGGGCTCTCAGGAG 758

LOCUS HSA314905 34667 bp DNA linear PRI 21-MAY-2002
DEFINITION Homo sapiens partial OBSCN gene for obscurin, exons 31-56.
ACCESSION AJ314905
VERSION AJ314905.1 GI:21104337
KEYWORDS OBSCN gene; obscurin.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1
AUTHORS Young, P., Ehler, E. and Gautel, M.
TITLE Obscurin, a giant sarcomeric Rho-GEF protein involved in sarcomere assembly
JOURNAL Unpublished
AUTHORS 2 (bases 1 to 34667)
TITLE Direct Submission
JOURNAL Submitted (22-MAY-2001) Gautel M.S., Physiologische Biochemie,
Max-Planck-Institut fuer molekulare Physiologie, Otto-Hahn-Strasse
11, Dortmund, 44227, GERMANY

FEATURES
source 1. 34667
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

gene
mRNA
CDS
join(247..519,1332..1598,1817..2082,4388..4664,5087..5353,6008..6274,6375..6638,6939..7202,10212..10475,10837..11100,11667..11930,12466..12729,13094..13357,16768..17031,17573..17836,18196..18459,21989..22252,22617..22880,23476..23700,24686..24952,25215..25481,25573..25846,26423..26697,27413..27609,27798..27873,34161..34433)
/gene="OBSCN"
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

protein 18="CAC85752.1"
/protein_18="CAC85752.1"
/db_xref="GI:21104338"
/translation="LHVGITRRLKTMTEVLEGECSFECVLSHESASDPAMTVGKTV
GSSRFQATROGRKYILVREAPSDAGEVSVRGLTSKSLIYREPAIIXPLED
OMVAPGDEVERCELSRAGTPVHMLKORKAIRKOKDVCEGTAMMIVIGASLKD
GEYTCVEYASKSTSLHYEERKANFTBELTTLQYBEKTAIVFTCTHPAATVYTRK
LLEIRASGKHQPSOGLTLRTTISALEAUSDYTCIDGQKSRQALVDQGRVHLL
EDLEDDVQVESSATFRCPISPAVPEVHMLDPTCDIANLEIDAPPGYHVLTLR
OLAKDKSGTIFYEAGDSALALVTVEPSFSELDLDTATTEGDEDLVLVETSCDI
PVCTAKQKTLIRGAGRCOLSHGRLAOLITGALTLOBSGRYKCEAGSSSIVVHA
RPAVFOBALKDLLEVBGAATLRCLVSVVAPVWCKVANKVLRBGDKSLTBOEKMEL
LVVNRNRPODSGRYSCSGDOTTSTATLTVALTPAOFICLKNKATBEATATLRBELS
KAAPVERKSGSETLRDGRVCLRDGACELQIGLAVNDAABYSVCGBERTSLSLT
IRPMPAHFIRGLRHQESIECVGERTSATLTVALPAKTEBGRNEBAVEGATLMACE
CELDICGLAVADAGEYSCVCGERTSATLTVALPAKTEBGRNEBAVEGATLMACE
LSKVAPEVRKGPENLRDGRVILROBERTSALQICGLAMADAGEYLCVCGERTSAT
LTITAPLPAFIEDVNOBARGATVALQCELSAAPVWRGSETLRDGRVSLRDG
TKCELOIRGLAMADAGEYSCVCGERTSATLTVALPAKTEBGRNEBAVEGATLMACE
CELSKAPVEMKKGHEITLRDGRVSLRDGACELQIGLAVNDAABYSVCGBERTS
AMLTVPAPSKFIEGLRNEATEBDTATLWCELSKAAPVERKKGHEITLRDGRVSLRD
DSCCELOIRGLAVADAGEYSCVCGERTSATLTVALPAFIEDVNOBARGATVAL
LOCELSKAPVEMKKGSETLRDGRVSLRDGACELQIGLAVNDAABYSVCGBERT
TSATLTVALPAFIEDVNOBARGATVALQCELSKAPVEMKKGSETLRDGRVSLRD
KODTRCELOIHDLSVADAGEYSCVCGERTSATLTVALPAFIEDVNOBARGATVAL
ATLQCELSKAPVEMKKGHEITLRDGRVSLRDGACELQIGLAVNDAABYSVCGBERT
ERTSATLTVALPAFIEDVNOBARGATVALQCELSKAPVEMKKGHEITLRDGRVSLRD
SLKODGTSCELOIRGLAVADAGEYSCVCGERTSATLTVALPAFIEDVNOBARGATVAL

CDS

```
/gene="OBSCN"  
72..19934  
/gene="OBSCN"  
/function="t1cin binding"  
/codon_start=1  
/evidence=experimental  
/product="Obscurin"  
/protein_id="CAC44768.1"  
/db_xref="GI:15026974"  
/db_xref="GOA:O96AA2"  
/db_xref="UniProt/TREMBL:O96AA2"  
/translation="MDQPSGARPRFLTRKAVVSVSGKDATISQIVGNPPTQVSW  
KDQDPYRAGARFLADQSDGLRYLTLLDIALSGQVYCRARNIIGAFVAVGQVAVAE  
EACABQPHFLNTPSIRVRESSEATFRCVGSPPRAVWSMDGRILGDPDRYAV  
EIEGASALAIRARPRDGGTYEVRARENPLGAASAAALVDSADADTLSPSTSTRA  
LMAHRRRREARMAEGAPASPSTGTRCTVTEGKARLSCTYVGEPEKPTKXQGO  
LVEGRHHVVYEDAEENFVLKIFCKSDRGILCTTASNLVQOTSSVTVVVRPAVP  
FKRLDILVEGRESATFLCEVPDSTEAMFEBRILMAKATGIEEBSTERLTVR  
NVASDDAVYICETPESGRVALA VQNLRLKLPKRTAVRGDTAMPCEVLAVGP  
VHMLRNOEVVAGGVAISAGTRHTLTI SQCLEEDVGVAFMAGDCQSTRCVSAP  
RKPLQPPVDPVYKAMNBSVILISMPPIGEPPTIDGLVEKKILGTYTMRCHA  
EMVATPELTVADVAEENGFQFVSLNSPFGSPYLEPPTVHLAPKLARTPLKAVQA  
VEGEVTFSDTLTVASAGWFLDGLAKASSVYEHICDRTHTLTIREVPASIHQOL  
KVFANGIESIRMEVRAARGLTANKPRAAARVTLRLHEBAQLALSDQAAVATWL  
KQRTISPEPKTYVQASGRVLDVADVADAGLVECSRGRGLAYQLSVGLARFL  
HKOMASCVDAVAGPAQOFECEITSEAHVHVHMKDMEILHSGERFLQEDVGRHLV  
AATVTRQDGGTVSCRVEGSDVDFRLVSEPKVFAKEOLARLKLQAEAGASATLSCEV  
AQOITEVTVKDKKLSSTSSKVMENATGCTRLVVOOAGQADAGEYSCAAGGRLSFH  
LVKEPKVFAKDOVAHSEVQAEAGANATLSCEVQAQOAEVMMYDQDKLSLRYH  
BAKGRRLVVOOAGKTDAGDYSCEARGORVSRHLITTPKRMFAKESVHNVQAE  
BAGAMLSCVNAQOITEVTVKDKKLSSTSSKVMENATGCTRLVLPQAKADAGEYSC  
EAGQDVSFHLHITTEPKVFAKESVHNVQAEAGTATLSCVAPQOITEVTVKDKK  
KLSSSKVMEVYKCTRLVVOOAGQADAGEYSCAAGGRLSFHITTEPKVFAKEQ  
LVHNEVRETAGASATLSCEVQAQOITEVTVKDKKLSSTSSKVMENATGCTRLV  
GOADAGEYTCGAAGGRLSFHLPVSEPKVFAKEOLARLKLQAEAGASATLSCEV  
TEVTVKDKKLSSTSSKVMENATGCTRLVVOOAGQADAGEYSCAAGGRLSFH  
EKKVFAKOPVHREVOQAAGSTLSCVNAQOITEVTVKDKKLSSTSSKVMENATG  
CTRLVVOOAGQADAGEYSCAAGGRLSFHLPVSEPKVFAKEOLARLKLQAEAGAS  
TLSCVQAQOITEVTVKDKKLSSTSSKVMENATGCTRLVVOOAGQADAGEYSCAAG  
OQLSFHLVAEPKVVFAKOPHAREVOAEAGASATLSCEVQAQOITEVTVKDKKLS  
SSKVEVFAKCTRLVVOOAGQADAGEYSCAAGGRLSFHLPVSEPKVFAKEQ  
PLVKEHEDIILTATLTPSAATVWLKDGVEVIRRSKHETASQDHTLTLYHQVQL  
DSATVSCRVAEGODPEPVQVEVAAKFCGLLEPVCGLEGGTTLTLCESLPACAVVR  
CQNTOPRVKRPQMAECPVRLTYVLGRABEAGEVCSRDHTSAQIVSPVAVVR  
PNSGLTVAAEAGEATPQCVVPSQDVAVFRDALLQPSKFAISQSGHSLTIS  
DLVLEDAQOITVEAGASSALRVREAFLVFKKLEPQTEVERSSVLEVELTTPM  
ELMTRNATALLAPEGKVEIHAAGRRHLVHNVPADRGFCGCTPDQTOAKLVEM  
ROVRLVRGLQAVAREOGGTATWVOLSHADVDSMTRGRLRQOGPTCHLAVRGMMT  
LTLGRLPDSGLAVPKAGVHTSARLVYTELPSVPSRLDQVYVTEKATTLBCELS  
RNVVVRMLKQDVEIRAGKTALIAQACRSLLTIRCEPADQVVRCAHQAQSSAV  
KYQKRTYTLIRVLAEDAGEIQVAENAESRAQLRYELPVLVLRPRDXIAKMKHR  
GVLEQVFAESAOVMFKGSOELQPGKTELVSDGLYRLLISDVHADDEPTVYDAG  
DVTECAOFPVEEOSITIVGLQDVTYMEBAPAMPFCETSIPEVRPKMLDKTYVAC  
GNVLEOEGTVHRLMLRRTCSITMGVTPHYTKVSSSARIVAVSDIPIVLTIRPLEKQ  
RELQSVLVCDPRPAKAVOMVYKDDTPIVSPSKPKMSLEGMAELITLILMPADAGV  
RCQASASSTTVTYEARVYTYTGPLQAEATEBMAWPSCELSHEDVEWMSLNGMP  
LYNDFHEISHKGRHTVLKYSIQADAGIVASSLKYSTARLEVRKVPVFKALD  
DLSAEERGTALAQCEVSDPEAVVRKDVQGPDKDPLHTAGTRLVAVHVPD  
AGLYTCHVGSSETRARVVDLHVGITRLKTMELVEBESGFEVLSHESASDPAW  
TVGKTVGSSSSRFQATROGRKYLIVREASRLDAGEVYVSRGLTSKSLIVRBPAA  
IIKPLEDOMVARGEDVELRCESLSPAGTSAVILKDKRAIKRQKIDVYVEGTMALIVR  
GASLKDAGEYTCVEPASSTASLHYEEKANCETELTNLYVEEKGTAQFOTCTHPAA  
TVMKRGLLEIRASGKHQPSQEGTLTLTISALEKADDTYTQIGQAQSOQLVQV  
RRVHILIEDLEVNDVQESSATFRCSIPANVEPVMFJDKPLPLANENEDADPGY  
HVLTLRLALKDVSQGITFEADQORASALRYEKSPVSRDELTPATITIEGDLTVCE  
TSTCDIPMCWTKDKGTLRGSARCOLSHGHGAOLLTGATLTDQSGRNVLCRBDKSLQ  
IVRVHARVPRFOBALKDLLEVLGGAAITRCTLSSVAAVVRKVCIGNLVRBDKSLQ  
EGAMLELYVRRLRQDSGRYSCSFGDQYTSATLTVTALPAFICGLNRKKEATECAT  
LRCELSKTAIVWRKGSITLNDGDIYCRQDAMCELOIRLAWDAEASVCSGBERL  
TSASLITRPMFAHFGRLRHQESIEGATPATRCELSTKAPVEMKGBESLSDGRHL  
RQDAVCELQICGLAVADAGEYSCVCGEERTSATLYVALALAKTEGIRAGEYLVCCO  
AMUCELSKAPVEMKGPENLRDGRYILIEOGRCELOICGLMADAGEYLVCCO  
ERTSATILTRALPARFIEDVKNQAREBAGATVLOCELSNAPVWRKGSITLRDGRY
```

Qy	1	GAAGACCAAGTGGGGTGGCGCCAGAGAGACGTGAGAGCTGTGAGCTGTCAACGGCG	60
Db	8571	GAAGACCAAGTGGGGTGGCGCCAGAGAGACGTGAGAGCTGTGAGCTGTCAACGGCG	8630
Qy	61	GGAACGCGCGTGGACGTGGCGTGAAGGACAGGAAGCCATCCGCAAGGCCAGAGTATGAT	120
Db	8631	GGAACGCGCGTGGACGTGGCGTGAAGGACAGGAAGCCATCCGCAAGGCCAGAGTATGAT	8690
Qy	121	GTGGTCTGCGAGGAGCAAGATGCGCATGCTGTGATCCGCGGGGCTCGCTCAAGACGCG	180
Db	8691	GTGGTCTGCGAGGAGCAAGATGCGCATGCTGTGATCCGCGGGGCTCGCTCAAGACGCG	8750
Qy	181	GGGAGGTACACGTGTGAGGTGAGGCGCTTCCAGAGACACAGCCCTCCATGTGGAAGAA	240

SLRDQGTKELOIRGLAMADNGEYSCVCGSESTSAMLTVALPIKFTREGIRMEATEG
ATAVRCELSTKAPVEMKGBESLSDGRHSIRROGARELOIRGLVADAGEYLCMC
GERTSAMLTVALPIKFTREGIRMEATEGATVATIMCELSKAPVEMKGBESLSDGR
RHSIRROGARELOIRGLVADAGEYSCVCGSESTSAMLTVALPIKFTREGIRMEATEG
RHSIRROGARELOIRGLVADAGEYSCVCGSESTSAMLTVALPIKFTREGIRMEATEG
BGAATVLOCELSKAPVEMKGBESLSDGRHSIRROGARELOIRGLVADAGEYSCVCG
VCGERTSAMLTVALPIKFTREGIRMEATEGATVATIMCELSKAPVEMKGBESLSDGR
ANGREESTRLOGCTAEVLQDLOREDTGTGTGCGSAGTATVLAAPVRLREIHO
EYDDEGTAHLCCELSTKAPVEMKGBESLSDGRHSIRROGARELOIRGLVADAGEY
DYTCDSHGTQMSASLSTVPRPKTRRLQSLSEQETGIDARLCCQLSDSAGVAVOMLK
BGEVILHAGPKTMRQGTARELILHOLEKADGEEACVYTGQHLKLSAPPHIAGEVTRCD
RELVAEVLVADAGEYSCVCGSESTSAMLTVALPIKFTREGIRMEATEGATVATIM
GVTLPEDJAGEYSCVCGSESTSAMLTVALPIKFTREGIRMEATEGATVATIMCELS
DEARWALGVPLOANEMNDITVEOQTLLHILHKTUPLDAGTVSPHVGCSSEADLKY
TAKNTVGLNEVLEBGEALFEQOLSPVAAATWMLDDEPTEGSMNAPVPRVENC
LRLHLLKNLAPDSCTPFLAGDMTSAFLTVRSMRLEIIEPLKMAVRAQAQRFT
CTLSRAVYGERBASTINKAAVQPDSDMTVALPDRSQHLKLSAPPHIAGEVTRCD
VAASARLVTLGIPDEDEAVASHSHTVLSMAPMDDGGGLGCVREVEAGATVQ
WMLCHEILVPEECVVDGLAPETVFAVVAQVAGAPVHLPTVRLAEPKVPVQ
PSAPESROVAAAGEYSCVCGSESTSAMLTVALPIKFTREGIRMEATEGATVATIM
FLAEDQGEYHGLAQGSI CPAATFOVALSPASVDEARQPSLPREAAQEGDHLIWEA
LARKRMRPEPTLISLTPEDRSGRLPOBAEVAPDLSGISTADELARTGDADL
SHTSDDSRAGTSLVTVLKKAGRGTSPLASVGAAPASVPRQOQEEPLAARVP
LGISTKOLGDPSPMDKAIVKIQAAFGKGVSPKASVGAAPASVPRQOQEEPLAARVP
LECVVASADVRAAMLKDGVELTDRGHHHIDQGTCSLLIAGLDRADACCTCOVS
NKFGVTHSACVAVVSGSESAESSGGELDQAFRAARLRLHRLPRTSSPAPVSEDLF
LSADGPAPE
VBOGPRVEMCISKETRAPVYPPPRPBLTSDAPVLTGLQNEVDDGTPVSDCV
VTQGPMSVPRFKKGLLEEDHMTINDQGGHQLITAVPDMGVYRCLANSWG
VSTYAEIRVLTSTDTYADADAGEYSCVCGSESTSAMLTVALPIKFTREGIRMEATEG
BLRDLQVAPGTRLAKFOLKGVYAPAPRYLWFKTSPKSPQKQWASPAVLDRLKLP
VTRDESGQYAAVYINAMGAAYASARLVAGVDEEKPASVHQLVPLMRLEETPK
KVKGSSITFVKEGEPVPTVHMLREAEAGVMIIGDTPGTVALVSAPOHSHVILD
VGRQHGQYTCI ANAAGOLCASLISLHSLAEPKREOKYKALISPTLOCTQTAISA
QPGOTASPADQCGORKEPPLAALKALHLSLISAESEDEFLQTSITTEMASCTIOA
KLQVPGSDSDSDSTPSAPPHGSRSSSIQESSSEEDDAGELFDYLVAVADYLV
PLGABQDITLREQVYVLDAAHPLMLWTKTKSPSPQKQWASPAVLDRLKLP
EMGAELAEPEFGEVDEYVAKRLSYIOELISEQAFVEELQGLSHHLOHLRCPH
VPIVAGQKAVIFENRNDIGRHSFLOELQOCCPTDVMACFIKNOAFPOCYVPLK
GRVQESVAVTAIOBFPYKKEFALLAGDSDPPDLOHMYLEPOVPRORVQALLK
LILNKARNRQNCLEOAYAVATLORANKLHSLMERYPTLOALPEIRQGH
IWEAGANRPNKGNHRHVTFLPNHVLICPRDESDTDTVSYFRMMKLSSITLND
QVSDDAFAPEWQREDSVRYLLQARTALIKSSWVEICQIQORLALPWRPDPFER
ELADCTAIEGTATVLAQVYTPKPSVWPKDGAQVDDPHLILIEPDSQALILDS
LTVGDSQALMFAASAAGNCSTGLKILVQVPRVNVVLASPFVEGDAOFTCTICGA
PYIRMYKQDALLITGNKFOITSEPSGLVIVTIRASKEIDGLTYCELVNRGSGAR
ASAPLRIOSMLAOBOCHREOLVAAIEVTEGKEVYKXTYVILIEITTTVYKSRGCR
RSPEKSSRSRPSRSASPLRGLIADPLVLPDAGQORPEAEFGQVPPVPTLVTA
EASHPALPGLSGPOPKAVEVEBTI EVAVKMKGPQGVSPTEVPRSSGHLFTLEGATP
GGDPSNNSNNKKLLOEAMAOAGTAMVGRVREMPVFRVARSVMAASGMSLEBETM
EEAGEEEDGEDDAFVTEESODTSLIGRDXKILTNHGRMLTADLEIYVGEGETFHC
GGPQGPADDPPECVSVIOREIGEPTVGGVLLSVGRALGPRDGLFRPEPQAFSP
GPOVRSLEGSFLLREAPRVPASAPMTGSCFTRIRRSADSGSFTTELSTQTVNFG
TVGETVTLHICPRDGBEAP"

Db 1541 GGCAGTACAGCTGTGAGTGGAGCTTCCAAGACAGCAGCCATCATGTGAGAGGT 1600
Qy 241 AAA 243
Db 1601 AAA 1603

RESULT 9
AL53593
LOCUS 135964 bp DNA linear PRI 13-FEB-2002
DEFINITION Human DNA sequence from clone RPS-1139B12 on chromosome 1q42.1-43,
complete sequence.
ACCESSION AL53593
VERSION AL53593.33 GI:18673899
KEYWORDS HTG:
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Almeida J.
REFERENCE 1 (bases 1 to 135964)
AUTHORS Direct Submission
TITLE Submitted (13-FEB-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Feb 14, 2002 this sequence version replaced gi:1797879.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Emi, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WormPEP; Information on the WormPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RPS-1139B12 is from the library RPCI-5 constructed by the group of
Pleier de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2
This sequence is the entire insert of clone RPS-1139B12.
Location/Qualifiers
1..135964
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="q42.1-43"
/clone="RPS-1139B12"
/clone_1b="RPCI-5"
25928..26054
/note="Sequence from overlapping clone RPI1-520H14
(AL535910). Assembly confirmed by restriction digest."

ORIGIN

Query Match 59.8%; Score 239.8; DB 9; Length 135964;
Best Local Similarity 99.2%; Pred No. 1e-38; 2; Indels 0; Gaps 0;
Matches 241; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
1 GAAGACCAAGTGGTGGCCGACGAGAGAGAGCTGAGCTGCTGAGCTGTACAGGCG 60
|||||

Db 70374 GAAGACCAAGTGGTGGCCGACGAGAGAGAGCTGAGCTGCTGAGCTGTACAGGCG 70433
Qy 61
Db 70434 GAAGACCCGTCAGCTGCTGAGAGAGAGAGGCTATCCGACAGAGCCAGAGTATAT 70493
Qy 121
Db 70494 GTGCTTCGAGAGGACGATGCGATGCTGATCCGCGGGGCTGCTCAAGACGCG 70553
Qy 181
Db 70554 GGCAGTACAGCTGTGAGTGGAGCTTCCAAGACAGCAGCCATTCATGTGAGAGA 240
Qy 241 AAA 243
Db 70614 AAA 70616

RESULT 10
AC02657 164766 bp DNA linear HTG 01-SEP-2000
LOCUS Homo sapiens chromosome 1 clone RPI1-245P10, WORKING DRAFT
DEFINITION SEQUENCE, 31 unordered pieces.
ACCESSION AC02657
VERSION AC02657.4 GI:9958202
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Waterston R.H.
REFERENCE 1 (bases 1 to 164766)
AUTHORS The sequence of Homo sapiens clone
TITLE Unpublished
JOURNAL 2 (bases 1 to 164766)
AUTHORS Waterston R.H.
TITLE Direct Submission
JOURNAL Submitted (22-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 1, 2000 this sequence version replaced gi:7637349.
COMMENT

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Project Information
Center project name: H NH0245P10
----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 139884 bases at least Q40
Consensus quality: 147686 bases at least Q30
Consensus quality: 151469 bases at least Q20
Insert size: 169000; agarose-fp
Insert size: 161074; sum-of-contigs
Quality coverage: 3.60 in Q20 bases; agarose-fp
Quality coverage: 3.92 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 1124: contig of 1124 bp in length
* 1125 1124: gap of unknown length
* 1126 1124: gap of unknown length

ATVILQCEI¹ESKA²PVEMRKGS³ETL⁴IBGD⁵YS⁶LRQD⁷GTCE⁸LQ⁹IGLA¹⁰VEDTGE¹¹YL¹²CV¹³
CG¹⁴RTIS¹⁵ETL¹⁶VP¹⁷LP¹⁸PF¹⁹LD²⁰NTD²¹BAR²²GA²³TAT²⁴IC²⁵LS²⁶VA²⁷PVEMRK²⁸GET²⁹TL³⁰RD³¹
RHS³²LRQD³³GTCE³⁴LQ³⁵IG³⁶LS³⁷VAD³⁸GEYS³⁹CV⁴⁰CG⁴¹ERT⁴²SA⁴³TL⁴⁴T⁴⁵RE⁴⁶AT⁴⁷EG⁴⁸AT⁴⁹AL⁵⁰QCEI⁵¹ES⁵²K⁵³
VAP⁵⁴VP⁵⁵EMRK⁵⁶GET⁵⁷TL⁵⁸RD⁵⁹DR⁶⁰YN⁶¹LK⁶²QD⁶³GTCE⁶⁴LQ⁶⁵IG⁶⁶LS⁶⁷VAD⁶⁸GEYS⁶⁹CV⁷⁰CG⁷¹ERT⁷²SA⁷³TL⁷⁴TV⁷⁵
KA⁷⁶OP⁷⁷ARE⁷⁸BLQ⁷⁹LA⁸⁰EG⁸¹ST⁸²AT⁸³LQCEI⁸⁴SE⁸⁵PA⁸⁶TVMS⁸⁷KG⁸⁸LO⁸⁹QANG⁹⁰RE⁹¹PR⁹²LRQ⁹³GT⁹⁴
AE⁹⁵VL⁹⁶OD⁹⁷LE⁹⁸D⁹⁹LTGE¹⁰⁰YTC¹⁰¹CG¹⁰²SA¹⁰³TS¹⁰⁴TL¹⁰⁵VA¹⁰⁶VR¹⁰⁷PL¹⁰⁸EL¹⁰⁹OH¹¹⁰ED¹¹¹EG¹¹²TAL¹¹³QCEI¹¹⁴
LS¹¹⁵GA¹¹⁶SVEM¹¹⁷RK¹¹⁸SLQ¹¹⁹LP¹²⁰CA¹²¹RY¹²²OW¹²³OG¹²⁴AA¹²⁵ELL¹²⁶VR¹²⁷CE¹²⁸QED¹²⁹AD¹³⁰GT¹³¹CD¹³²GT¹³³GH¹³⁴OSMA¹³⁵
SL¹³⁶VR¹³⁷PR¹³⁸PE¹³⁹K¹⁴⁰TR¹⁴¹LOS¹⁴²LE¹⁴³Q¹⁴⁴ET¹⁴⁵DI¹⁴⁶AR¹⁴⁷LL¹⁴⁸CCQ¹⁴⁹LS¹⁵⁰D¹⁵¹BS¹⁵²GA¹⁵³V¹⁵⁴Q¹⁵⁵ML¹⁵⁶EG¹⁵⁷EL¹⁵⁸LAG¹⁵⁹CT¹⁶⁰EM¹⁶¹
RSG¹⁶²AT¹⁶³RE¹⁶⁴LI¹⁶⁵HO¹⁶⁶LE¹⁶⁷K¹⁶⁸D¹⁶⁹GTGE¹⁷⁰YAC¹⁷¹YG¹⁷²Q¹⁷³K¹⁷⁴TAS¹⁷⁵LR¹⁷⁶VT¹⁷⁷GE¹⁷⁸PE¹⁷⁹TV¹⁸⁰IG¹⁸¹LV¹⁸²GL¹⁸³VAD¹⁸⁴ET¹⁸⁵DE¹⁸⁶
DV¹⁸⁷FS¹⁸⁸CE¹⁸⁹VS¹⁹⁰AG¹⁹¹AT¹⁹²GV¹⁹³Q¹⁹⁴CLQ¹⁹⁵PLQ¹⁹⁶SN¹⁹⁷EV¹⁹⁸TE¹⁹⁹VA²⁰⁰RR²⁰¹DR²⁰²IHT²⁰³LY²⁰⁴LK²⁰⁵GT²⁰⁶ED²⁰⁷AG²⁰⁸TV²⁰⁹SF²¹⁰
H²¹¹LN²¹²MS²¹³AD²¹⁴Q²¹⁵TV²¹⁶ARE²¹⁷VL²¹⁸LE²¹⁹PL²²⁰DV²²¹LS²²²EG²²³OD²²⁴AS²²⁵FC²²⁶RL²²⁷SR²²⁸AS²²⁹GE²³⁰EA²³¹ML²³²AG²³³VL²³⁴
Q²³⁵AN²³⁶EM²³⁷DI²³⁸VE²³⁹Q²⁴⁰GT²⁴¹L²⁴²IL²⁴³TL²⁴⁴IK²⁴⁵VT²⁴⁶LE²⁴⁷D²⁴⁸AG²⁴⁹TV²⁵⁰FS²⁵¹H²⁵²GT²⁵³CS²⁵⁴EA²⁵⁵OL²⁵⁶KA²⁵⁷XT²⁵⁸VA²⁵⁹IG²⁶⁰VL²⁶¹
V²⁶²LE²⁶³EG²⁶⁴EL²⁶⁵FE²⁶⁶Q²⁶⁷SO²⁶⁸PE²⁶⁹VA²⁷⁰HT²⁷¹WL²⁷²DD²⁷³EP²⁷⁴HT²⁷⁵SEN²⁷⁶AE²⁷⁷V²⁷⁸FF²⁷⁹ENG²⁸⁰LR²⁸¹HL²⁸²LL²⁸³KN²⁸⁴LR²⁸⁵
OD²⁸⁶CR²⁸⁷VT²⁸⁸PL²⁸⁹AG²⁹⁰MT²⁹¹SA²⁹²PL²⁹³TV²⁹⁴A²⁹⁵"
247. .519
/gene="OBSCN"
/number=31
520. .1331
/gene="OBSCN"
/number=31
1332. .1598
/gene="OBSCN"
/number=32
1599. .1816
/gene="OBSCN"
/number=32
1817. .2082
/gene="OBSCN"
/number=33
2083. .4387
/gene="OBSCN"
/number=33
4388. .4664
/gene="OBSCN"
/number=34
4665. .5086
/gene="OBSCN"
/number=34
5087. .5353
/gene="OBSCN"
/number=35
5354. .6007
/gene="OBSCN"
/number=35
6008. .6274
/gene="OBSCN"
/number=36
6275. .6374
/gene="OBSCN"
/number=36
6375. .6638
/gene="OBSCN"
/number=37
6639. .6938
/gene="OBSCN"
/number=37
6939. .7202
/gene="OBSCN"
/number=38
7203. .10211
/gene="OBSCN"
/number=38
10212. .10475
/gene="OBSCN"
/number=39
10476. .10836
/gene="OBSCN"
/number=39
10837. .11100
/gene="OBSCN"
/number=40
11101. .11666
/gene="OBSCN"

Query Match	59.8%	Score 239.8	DB 9	Length 34667
Best Local Similarity	99.2%	Pred. No. 1.3e-38		
Matches 241	Conservative	0	Mismatches 2	Indels 0
			Gaps	0
exon	/number=40 11667. .11930 /gene="OBSCN"			
intron	/number=41 11931. .12465 /gene="OBSCN"			
exon	/number=41 12466. .12729 /gene="OBSCN"			
intron	/number=42 12730. .13093 /gene="OBSCN"			
exon	/number=42 13094. .13357 /gene="OBSCN"			
intron	/number=43 13358. .16767 /gene="OBSCN"			
exon	/number=43 16768. .17031 /gene="OBSCN"			
intron	/number=44 17032. .17572 /gene="OBSCN"			
exon	/number=44 17573. .17836 /gene="OBSCN"			
intron	/number=45 17837. .18195 /gene="OBSCN"			
exon	/number=45 18196. .18459 /gene="OBSCN"			
intron	/number=46 18460. .21988 /gene="OBSCN"			
exon	/number=46 21989. .22252 /gene="OBSCN"			
intron	/number=47 22253. .22616 /gene="OBSCN"			
exon	/number=47 22617. .22880 /gene="OBSCN"			
intron	/number=48 22881. .23475 /gene="OBSCN"			
exon	/number=48 23476. .23700 /gene="OBSCN"			
intron	/number=49 23701. .24685 /gene="OBSCN"			

RTSALTRPMPAHFICRLRHQESIEGATATLRLCELSKAAAPVWVRKRESLRDGRHS
LRQDGAECLEIQICGLAVADAGEYSCVCEERTSATLTVALPAKFTGLRNBEEAVEGA
TAMLCSELSKVPVWVRKGPENLRDGRYILRQGTRECELIQICGLAMADAGEYLCVCG
QERTSATLTTRALPARFEDVKNQEGATAVLQCELSAAPVWVRKSGSETLRDGR
YSLRQDGTCKCLOIRGLAMADTGBYSVCVGOERTSAMLTVRALPIKFTGELNBEATE
GATVALCELSKMAAPVWVRKGHETLRDGRHSRLRQDGRARCELOIRGLVAEDAGEYLCW
CGKERTSAMLTVRALPKFIEGLRNEATEGDTATLWCELSKAAAPVWVRKGHETLRDGR
DRLRQDGRSCELIQIRGLAVDAGEYSCVCEERTSATLTVALPAKFTGLRNBEEAVEGA
REGATVALCELSKAAAPVWVRKGPENLRDGRYILRQGTRECELIQICGLAMADAGEYLCVCG
QERTSATLTTRALPARFEDVKNQEGATAVLQCELSAAPVWVRKSGSETLRDGR
DGRYSLRQDGTCKCLOIRGLAMADTGBYSVCVGOERTSAMLTVRALPIKFTGELNBEATE
EMEGATATLQCELSKMAAPVWVRKGLRDLRDKYSLRQDGAVALCELOIRGLHGLAMADNGV
YSCVCEERTSATLTTRALPARFEDVKNQEGATAVLQCELSKAAAPVWVRKGPEN
LKQDGRYSLRQDGTCKCLOIRGLAMADTGBYSVCVGOERTSAMLTVRALPAKFTGLRNBEEAVEGA
NHEGATATLQCELSKMAAPVWVRKGPENLRDGRYILRQGTRECELIQICGLAMADAGEYLCVCG
QERTSATLTTRALPARFEDVKNQEGATAVLQCELSAAPVWVRKSGSETLRDGR
ETLRDGRHSRLRQENLRNPNPGGCGSELGSCSPAWTAKLDFIS"

ORIGIN

Query Match 100.0%; Score 401; DB 9; Length 4789;
Best Local Similarity 100.0%; Pred. No. 6.3e-71;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAGACCACTGGTGGCGCCAGGAGGACCTGAGGCTGCCTGTGAGCTCTCACGGGCG 60
Dy |||||
Qy 327 GAAGACCACTGGTGGCGCCAGGAGGACCTGAGGCTGCCTGTGAGCTCTCACGGGCG 386
Dy |||||
Qy 61 GGAACCGCCGCTGCACTGGCTGAAGACAGAGAGGACCTCCGCAAGAGCCAGAGATGATGAT 120
Dy |||||
Qy 387 GGAACCGCCGCTGCACTGGCTGAAGACAGAGAGGACCTCCGCAAGAGCCAGAGATGATGAT 446
Dy |||||
Qy 121 GTGTCTGCGAGGACGATGGCCATGTGTCTATCCGCGGGGCTCGCTCAAGAGACGCG 180
Dy |||||
Qy 447 GTGTCTGCGAGGACGATGGCCATGTGTCTATCCGCGGGGCTCGCTCAAGAGACGCG 506
Dy |||||
Qy 181 GGCGAGTACAGTGTGAGGTGAGGCTTCCAGAGCAGAGCAGGACCTCCATGTGGAAGAA 240
Dy |||||
Qy 507 GGCGAGTACAGTGTGAGGTGAGGCTTCCAGAGCAGAGCAGGACCTCCATGTGGAAGAA 566
Dy |||||
Qy 241 AAAGCAAACTGCTTCCAGAGGAGCTGACCAATCTGAGGTGGAGGAGGAGGACACAGCT 300
Dy |||||
Qy 567 AAAGCAAACTGCTTCCAGAGGAGCTGACCAATCTGAGGTGGAGGAGGAGGACACAGCT 626
Dy |||||
Qy 301 GTGTTCAGTGCAGACGAGGACACCCCGCGCCACAGTGACCTGCGCGCAAGGGGCTCTTG 360
Dy |||||
Qy 627 GTGTTCAGTGCAGACGAGGACACCCCGCGCCACAGTGACCTGCGCGCAAGGGGCTCTTG 686
Dy |||||
Qy 361 GAGCTACGGGCTCAGGAAGCAGCAGCCAGCCAGCCAGGAGGG 401
Dy |||||
Qy 687 GAGCTACGGGCTCAGGAAGCAGCAGCCAGCCAGCCAGGAGGG 727
Dy |||||

RESULT 5
LOCUS CQ730656 18524 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 16590 from Patent WO02068579.
ACCESSION CQ730656
VERSION CQ730656.1 GI:42305092

KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kifs, such as nucleic acid arrays, comprising a majority of
human exons or transcripts, for detecting expression and other uses
thereof

JOURNAL Patent: WO 02068579-A 16590 06-SEP-2002;
PE Corporation (NY) (US)

FEATURES
Location/Qualifiers
1..18524
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 401; DB 6; Length 18524;
Best Local Similarity 100.0%; Pred. No. 5.1e-71;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAGACCACTGGTGGCGCCAGGAGGACCTGAGGCTGCCTGTGAGCTCTCACGGGCG 60
Dy |||||
Qy 7170 GAAGACCACTGGTGGCGCCAGGAGGACCTGAGGCTGCCTGTGAGCTCTCACGGGCG 7229
Dy |||||
Qy 61 GGAACCGCCGCTGCACTGGCTGAAGACAGAGAGGACCTCCGCAAGAGCCAGAGATGATGAT 120
Dy |||||
Qy 7230 GGAACCGCCGCTGCACTGGCTGAAGACAGAGAGGACCTCCGCAAGAGCCAGAGATGATGAT 7289
Dy |||||
Qy 121 GTGTCTGCGAGGACGATGGCCATGTGTCTATCCGCGGGGCTCGCTCAAGAGACGCG 180
Dy |||||
Qy 7290 GTGTCTGCGAGGACGATGGCCATGTGTCTATCCGCGGGGCTCGCTCAAGAGACGCG 7349
Dy |||||
Qy 181 GGCGAGTACAGTGTGAGGTGAGGCTTCCAGAGCAGAGCAGGACCTCCATGTGGAAGAA 240
Dy |||||
Qy 7350 GGCGAGTACAGTGTGAGGTGAGGCTTCCAGAGCAGAGCAGGACCTCCATGTGGAAGAA 7409
Dy |||||
Qy 241 AAAGCAAACTGCTTCCAGAGGAGCTGACCAATCTGAGGTGGAGGAGGAGGACACAGCT 300
Dy |||||
Qy 7410 AAAGCAAACTGCTTCCAGAGGAGCTGACCAATCTGAGGTGGAGGAGGAGGACACAGCT 7469
Dy |||||
Qy 301 GTGTTCAGTGCAGACGAGGACACCCCGCGCCACAGTGACCTGCGCGCAAGGGGCTCTTG 360
Dy |||||
Qy 7470 GTGTTCAGTGCAGACGAGGACACCCCGCGCCACAGTGACCTGCGCGCAAGGGGCTCTTG 7529
Dy |||||
Qy 361 GAGCTACGGGCTCAGGAAGCAGCAGCCAGCCAGCCAGGAGGG 401
Dy |||||
Qy 7530 GAGCTACGGGCTCAGGAAGCAGCAGCCAGCCAGCCAGGAGGG 7570
Dy |||||

RESULT 6

HSJ2535

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

FEATURES

Location/Qualifiers

1..20435

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="1"

/map="1q42.1"

/cell_type="cardiomyocyte"

/tissue_type="cardiac"

/dev stage="adult"

1..20435

gene

HSJ2535 20435 bp mRNA linear PRI 14-SEP-2001
Homo sapiens mRNA for obscurin (OBSCN gene).

AJ002535

AJ002535.1 GI:15026973

OBSCN gene; obscurin.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1

Young, P., Ehler, E. and Gautel, M.

Obscurin, a giant sarcomeric Rho guanine nucleotide exchange factor

protein involved in sarcomere assembly

J. Cell Biol. 154 (1), 123-136 (2001)

21342081

11448995

2 (bases 1 to 20435)

Gautel, M.S.

Direct Submission

Submitted (31-OCT-1997) Gautel M.S., Structural Biology Division,

European Molecular Biology Laboratory, Meyerhofstr. 1, Heidelberg,

69117, GERMANY

Revised by author 20-JUL-2001

Location/Qualifiers

1..20435

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/chromosome="1"

/map="1q42.1"

/cell_type="cardiomyocyte"

/tissue_type="cardiac"

/dev stage="adult"

1..20435

TITLE
JOURNAL

REMARK
COMMENT

Direct Submission
Submitted (10-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapps@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nih.gov
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAL Plate: 44 Row: c Column: 8.
Location/Qualifiers
1. .2761
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3845393"
/tissue_type="Placenta, choriocarcinoma"
/clone_lib="NIH MGC 21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"
<1. .2306
/codon_start=3
/product="Similar to titin"
/protein_id="AAH43916.1"
/db_xref="GI:27682000"
/translation="DNVRIQGRYTYLIRVLAEDAGETQFVAENAESRAQLRVKE
LPVTLRPLRKIAEMKHGRVLEQVSASQVRFKSGLOPQPKYELVSDGLYRK
LIISDVHAEDDTYTCADGDVYKSAQFFVEEQSITVIRGLQDVMPAPAFWPCETS
IPSVRPKWLLEGTVLAAGNVGLEQGTVHRLMRCTCTMTGTVHFTVSKRSAR
LVSDIPVVLTPLEKTRGLQSVLSCDFRPAPKAVQYKDDTLPSPSKPWSLE
QMAELRLRLMPADAGVYRCAQSAHSSTVTEAREVTVTGLQDAEATEGWASF
SCELSHDEEVEWLSNGNPLVNDSPFHEISHKRRHTLVLSIQRADAGIVRASLSKVS
TSARLEVRKPVFLKALDDUSABERGTLAQCEVSDPEAHVVMKDGVLQSPDKDY
FLHTAGTRGLVHDVSDPADGLYTCVHGSSETRARVRVHDLHVGITKRLKTMVELEGE
SCSEFVLSHESASDPAMWTGGKTGSSSRFOATROGRKYILVVRAPSDADGEVVF
SVRGLTSAKSLIVRERPAALIKPLEDOWAPGEDVELRCELSTRAGTPVHMLKORKAIR
KSQYDVVCEGTAMLVTRGASLADAGEYTCSEVASKSTASLHVEKANCFTBELTNL
QVEKGTAVFTCKTEHPAATVWRKGLLELRASGHQPSQEGTLRLITLSALEKADSD
TYTCDIGQAQSRALLVQGEAAKCARA"

FEATURES
source

CDS

ORIGIN

Query Match 100.0%; Score 401; DB 9; Length 2761;
Best Local Similarity 100.0%; Pred. No. 6.9e-71;
Matches 401; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GAAGACCAAGTGGTGGCCAGGGAGGACCTGGAGCTGCCTGTGAGCTGTCCACGGCG 60
|||||
Db 1773 GAAGACCAAGTGGTGGCCAGGGAGGACCTGGAGCTGCCTGTGAGCTGTCCACGGCG 1832
|||||

Qy 61 GGAAGCCCGTGCCTGCTGAAGACGACGAGGCCATCCGACAGACGACGAATATGAT 120
|||||

Db 1833 GGAAGCCCGTGCCTGCTGAAGACGACGAGGCCATCCGACAGACGACGAATATGAT 1892
|||||

Qy 121 GTGCTCTGCGAGGCGACGATGCCATGCTGCTCATCCGCGGGGCTCGCTCAAGGACGCG 180
|||||

1893 GTGCTCTGCGAGGCGACGATGCCATGCTGCTCATCCGCGGGGCTCGCTCAAGGACGCG 1952
181 GCGCAGTACACGCTGTGAGGTGGAGCTTCCAAAGACACAGCCAGCCTCCATGTGGAAGA 240
|||||
1953 GCGCAGTACACGCTGTGAGGTGGAGCTTCCAAAGACACAGCCAGCCTCCATGTGGAAGA 2012
|||||

241 AAAGCAAACTGCTTCACAGAGAGCTGACCAATCTGCAGGTGGAGGAGAAAGGCACAGCT 300
|||||
2013 AAAGCAAACTGCTTCACAGAGAGCTGACCAATCTGCAGGTGGAGGAGAAAGGCACAGCT 2072
|||||

301 GTGTTCACGTGCAGAGACGAGACCCCGGGGACACAGTGCCTGGCGAAGGCGCTCTTG 360
|||||

2073 GTGTTCACGTGCAGAGACGAGACCCCGGGGACACAGTGCCTGGCGAAGGCGCTCTTG 2132
|||||

361 GAGCTACCGGGCTCAGGGAAGCACCAGCCAGCCAGCGAGGAGG 401
|||||

2133 GAGCTACCGGGCTCAGGGAAGCACCAGCCAGCCAGCGAGGAGG 2173
|||||

AB046776 Homo sapiens mRNA for KIAA1556 protein, partial cds.
LOCUS AB046776.1 GI:10047176
DEFINITION Homo sapiens mRNA for KIAA1556 protein, partial cds.
ACCESSION AB046776
VERSION AB046776.1
KEYWORDS Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Nagase, T., Kikuno, R., Nakayama, M., Hirose, M. and Ohara, O.
Prediction of the coding sequences of unidentified human genes.
XVIII. The complete sequences of 100 new cDNA clones from brain
which code for large proteins in vitro
DNA Res. 7 (4), 273-281 (2000)
20450683
PUBMED 10997877
REFERENCE 2 (bases 1 to 4789)
Ohara, O., Nagase, T. and Kikuno, R.
Direct Submission
Submitted (03-AUG-2000) Osamu Ohara, Kazusa DNA Research Institute,
Department of Human Gene Research; 1532-3, Yana, Kisarazu, Chiba
292-0812, Japan (E-mail: cdnaif@kazusa.or.jp;
URL: <http://www.kazusa.or.jp/huge>, Tel:81-438-52-3913,
Fax:81-438-52-3914)
Location/Qualifiers
1. .4789
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="fh18619"
/notes="vector:pBluescriptII SK plus"
1. .4789
/gene="KIAA1556"
<1. .>4789
/gene="KIAA1556"
/note="Start codon is not identified."
/codon_start=3
/product="KIAA1556 protein"
/translation="BABI3382.1"
/db_xref="GI:10047177"
/translation="RARVRVHDLHVGITKRLKTMVELEGSCSEFVLSHESASDPAM
WTGGKTGSSSRFOATROGRKYILVVRAPSDADGEVVFVSRGLTSAKSLIVRERPA
ALIKPLEDOWAPGEDVELRCELSTRAGTPVHMLKORKAIRKSQYDVVCEGTAMLVTR
GASLADAGEYTCSEVASKSTASLHVEKANCFTBELTNLQVEKGTAVFTCKTEHPA
ATVWRKGLLELRASGHQPSQEGTLRLITLSALEKADSDTYTCDIGQAQSRALLVQGE
AAKCARA"